



Universidad de San Carlos de Guatemala
Facultad de Ingeniería
Escuela de Ingeniería en Ciencias y Sistemas

**EVALUACIÓN TÉCNICO ECONÓMICA DE LA
IMPLEMENTACIÓN DE UN SISTEMA DE SERVICIO
VIRTUAL EN EL MEDIO GUATEMALTECO.**

Alain Orión Reyes Bojorquez

Asesorado por el Ing. Jorge Armin Mazariegos Rabanales

Guatemala, marzo de 2010

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA



FACULTAD DE INGENIERÍA

**EVALUACIÓN TÉCNICO ECONÓMICA DE LA
IMPLEMENTACIÓN DE UN SISTEMA DE SERVICIO
VIRTUAL EN EL MEDIO GUATEMALTECO.**

TRABAJO DE GRADUACIÓN

PRESENTADO A LA JUNTA DIRECTIVA DE LA
FACULTAD DE INGENIERÍA

POR:

ALAIN ORION REYES BOJORQUEZ

ASESORADO POR EL ING. JORGE ARMIN MAZARIEGOS RABANALES

AL CONFERÍRSELE EL TÍTULO DE

INGENIERO EN CIENCIAS Y SISTEMAS

GUATEMALA, MARZO DE 2010

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
FACULTAD DE INGENIERÍA



NÓMINA DE JUNTA DIRECTIVA

DECANO	Ing. Murphy Olympo Paiz Recinos
VOCAL I	Inga. Glenda Patricia García Soria
VOCAL II	Inga. Alba Maritza Guerrero Spínola de López
VOCAL III	Ing. Miguel Ángel Dávila Calderón
VOCAL IV	Br. Luis Pedro Ortiz de Leon
VOCAL V	Br. Jose Alfredo Ortiz Herincx
SECRETARIA	Inga. Marcia Ivónne Véliz Vargas

TRIBUNAL QUE PRACTICÓ EL EXAMEN GENERAL PRIVADO

DECANO	Ing. Murphy Olympo Paiz Recinos
EXAMINADOR	Inga. Floriza Avila Pesquera
EXAMINADOR	Ing. Sonia Yolanda Castañeda Ramirez
EXAMINADOR	Ing. Marlon Antonio Pérez Turk
SECRETARIA	Inga. Marcia Ivónne Véliz Vargas

HONORABLE TRIBUNAL EXAMINADOR

Cumpliendo con los aspectos que establece la ley de la Universidad de San Carlos de Guatemala, presento a su consideración mi trabajo de graduación titulado:

EVALUACIÓN TÉCNICO ECONÓMICA DE LA IMPLEMENTACIÓN DE UN SISTEMA DE SERVICIO VIRTUAL EN EL MEDIO GUATEMALTECO,

tema que me fue asignado por la Dirección de la Escuela de Ingeniería en Ciencias y Sistemas, en junio de 2007.

A handwritten signature in black ink, consisting of a large, stylized initial 'A' followed by a series of horizontal and vertical strokes, and ending with a long, sweeping tail.

Alain Orion Reyes Bojorquez



Universidad de San Carlos de Guatemala
Facultad de Ingeniería
Escuela de ciencias y Sistemas

Guatemala, 2 de agosto de 2008

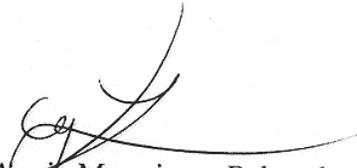
Ingeniero
Carlos Alfredo Azurdía Morales
Coordinador del Área de Trabajos de Graduación

Respetable Ingeniero Azurdia:

Por este medio le informo que como asesor del trabajo de graduación del estudiante universitario de la carrera de Ingeniería en Ciencias y Sistemas, Alain Orión Reyes Bojorquez, carne 200011398, he revisado y a mi criterio el mismo cumple con los objetivos propuestos para su desarrollo, según el protocolo de trabajo de graduación titulado "Evaluación Técnico Económica de la implementación de un sistema de Servicio Virtual en el medio Guatemalteco".

Agradeciendo su atención a la presente,

Atentamente,


Ing. Jorge Arrán Mazariegos Rabanales
Escuela de Ciencias y Sistemas
Asesor de trabajo de graduación
Colegiado: 5547



Universidad San Carlos de Guatemala
Facultad de Ingeniería
Escuela de Ingeniería en Ciencias y Sistemas

Guatemala, 20 de Agosto de 2008

Ingeniero
Marlon Antonio Pérez Turk
Director de la Escuela de Ingeniería
En Ciencias y Sistemas

Respetable Ingeniero Pérez:

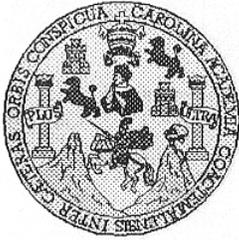
Por este medio hago de su conocimiento que he revisado el trabajo de graduación del estudiante **ALAIN ORION REYES BOJORQUEZ**, titulado: **"EVALUACION TECNICO ECONOMICA DE LA IMPLEMENTACION DE UN SISTEMA DE SERVICIO VIRTUAL EN EL MEDIO GUATEMALTECO"**, y a mi criterio el mismo cumple con los objetivos propuestos para su desarrollo, según el protocolo.

Al agradecer su atención a la presente, aprovecho la oportunidad para suscribirme,

Atentamente,


Ing. Carlos Alfredo Azurdia
Coordinador de Privados
y Revisión de Trabajos de Graduación

A circular stamp from the Escuela de Ingeniería en Ciencias y Sistemas, Universidad San Carlos de Guatemala. The text around the perimeter reads "UNIVERSIDAD DE SAN CARLOS DE GUATEMALA" and "ESCUELA DE INGENIERIA EN CIENCIAS Y SISTEMAS". The center of the stamp contains the text "COORDINACION TESIS".



Universidad San Carlos de Guatemala
Facultad de Ingeniería
Escuela de Ingeniería en Ciencias y Sistemas

Guatemala, 20 de Agosto de 2008

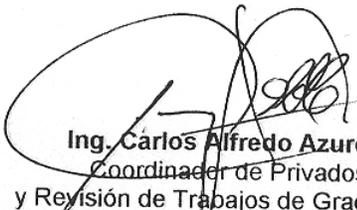
Ingeniero
Marlon Antonio Pérez Turk
Director de la Escuela de Ingeniería
En Ciencias y Sistemas

Respetable Ingeniero Pérez:

Por este medio hago de su conocimiento que he revisado el trabajo de graduación del estudiante **ALAIN ORION REYES BOJORQUEZ**, titulado: **"EVALUACION TECNICO ECONOMICA DE LA IMPLEMENTACION DE UN SISTEMA DE SERVICIO VIRTUAL EN EL MEDIO GUATEMALTECO"**, y a mi criterio el mismo cumple con los objetivos propuestos para su desarrollo, según el protocolo.

Al agradecer su atención a la presente, aprovecho la oportunidad para suscribirme,

Atentamente,


Ing. Carlos Alfredo Azurdia
Coordinador de Privados
y Revisión de Trabajos de Graduación



E
S
C
U
E
L
A
D
E
C
I
E
N
C
I
A
S
Y
S
I
S
T
E
M
A
S

UNIVERSIDAD DE SAN CARLOS
DE GUATEMALA



FACULTAD DE INGENIERÍA
ESCUELA DE CIENCIAS Y SISTEMAS
TEL: 24767644

El Director de la Escuela de Ingeniería en Ciencias y Sistemas de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer el dictamen del asesor con el visto bueno del revisor y del Licenciado en Letras, de trabajo de graduación titulado "EVALUACIÓN TÉCNICO ECONÓMICA DE LA IMPLEMENTACIÓN DE UN SISTEMA DE SERVICIO VIRTUAL EN EL MEDIO GUATEMALTECO", presentado por el estudiante ALAIN ORION REYES BOJORQUEZ, aprueba el presente trabajo y solicita la autorización del mismo.

"ID Y ENSEÑAD A TODOS"



Ing. ~~Marlon Antonio~~ Perez Turk
Director, Escuela de Ingeniería en Ciencias y Sistemas

Guatemala, 12 de marzo 2010

Universidad de San Carlos
de Guatemala



Facultad de Ingeniería
Decanato

Ref. DTG.091.2010

El Decano de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer la aprobación por parte del Director de la Escuela de Ingeniería en Ciencias y Sistemas, al trabajo de graduación titulado: **EVALUACION TÉCNICO ECONÓMICA DE LA IMPLEMENTACIÓN DE UN SISTEMA DE SERVICIO VIRTUAL EN EL MEDIO GUATEMALTECO**, presentado por el estudiantes universitarios **Alain Orion Reyes Bojorquez**, autoriza la impresión del mismo.

IMPRÍMASE.

Ing. Murphy Olympto Paiz Recinos
DECANO

Guatemala, marzo 2010

/cc
c.c. archivo.

ACTO QUE DEDICO A:

La Universidad San Carlos de Guatemala, que me ha dado la oportunidad de redescubrir el mundo con otros ojos.

Mi Madre quien me apoyo en todo momento, este esfuerzo es tanto suyo como mío, la llevo siempre en mi alma, pensamiento y corazón, gracias.

Mi familia que ha estado conmigo.

El Ingeniero Marlon Pérez Türk, su visión y trabajo ha hecho de la Escuela de Ciencias y Sistemas un mejor lugar.

El Ingeniero Armin Mazariegos, por asesorar el trabajo de graduación, por su confianza y apoyo, admiro su pasión por la enseñanza, he aprendido mucho de Usted.

ÍNDICE GENERAL

ÍNDICE DE ILUSTRACIONES	V
GLOSARIO	IX
RESUMEN	XI
OBJETIVOS	XIII
LÍMITES	XV
INTRODUCCIÓN	XVII
1 INNOVACIÓN	1
1.1 Innovación de soluciones: Máquina virtual + Linux minimal = Servicio Virtual	10
2 VIRTUALIZACIÓN	11
2.1 Qué es la virtualización	11
2.2 Historia de la virtualización	12
2.3 Cómo funciona la virtualización	15
2.4 Qué es una máquina virtual	17
2.5 Ventajas de una máquina virtual	18
2.5.1 Compatibilidad	18
2.5.2 Independencia	18
2.5.3 Encapsulación	19

2.5.4	Independencia del hardware	19
2.6	Ventajas de la virtualización	19
2.7	Problemas de la virtualización	20
2.7.1	CPU	21
2.7.2	Acceso directo a memoria DMA	22
2.7.3	Optimizaciones de administración de memoria	23
2.7.4	Acceso a red	24
2.8	Comparativa de distintos productos de virtualización	26
2.9	Herramienta de virtualización seleccionada: Vmware	27
3	LINUX	31
3.1	Qué es Linux?	32
3.2	Qué es una distribución de Linux?	33
3.3	TTYLinux	36
3.4	Sistema Alain	37
3.5	Modificación y compilación de Sistema Alain	39
4	RENDIMIENTO DE SISTEMA ALAIN EN UN AMBIENTE VIRTUAL	43
4.1	Equipo de prueba: Máquina de escritorio	44
4.1.1	Análisis del rendimiento de uso de CPU	45
4.1.2	Predicción del rendimiento de CPU para los distintos sistemas	50
4.1.3	Análisis del rendimiento de almacenamiento en disco	55
4.1.4	Predicción del rendimiento de disco para los distintos sistemas	60
	CONCLUSIONES	69

RECOMENDACIONES	73
REFERENCIAS ELECTRÓNICAS	75
BIBLIOGRAFÍA	81

ÍNDICE DE ILUSTRACIONES

FIGURAS

1. Iphone de Apple	3
2. Ipod de Apple	4
3. Mouse de una Imac	5
4. Disco duro sólido	6
5. One laptop per children	7
6. Arquitectura de aplicaciones en ambientes no virtualizados vs. ambientes virtualizados	15
7. Arquitecturas de virtualización	16
8. Windows XP ejecutándose en una máquina virtual en Slackware Linux 12	28
9. Tiempo de cálculo de PI para el sistema anfitrión, utilizando Slackware 12	46
10. Tiempo de cálculo de PI para Sistema Alain, corriendo en un sistema anfitrión, utilizando Slackware 12	47
11. Tiempo de cálculo de PI para Sistema Alain corriendo en un sistema anfitrión, utilizando Slackware 12	48
12. Diferencia de rendimiento para aplicaciones intensivas de CPU entre el Sistema Alaín y el sistema anfitrión.	49
13. Tiempo de descompresión para el sistema anfitrión, utilizando Slackware 12	56

14. Tiempo promedio descompresión para un Sistema Alain sobre un sistema anfitrión, utilizando Slackware 12	57
15. Tiempo promedio descompresión para un Sistema Alain sobre un sistema anfitrión, utilizando Windows XP SP2	58
16. Diferencia de rendimiento de acceso a disco entre el Sistema Alain y el sistema anfitrión	59

TABLAS

I. Comparativa de productos de virtualizacion	26
II. Paquetes del Sistema Alain	39
III. Tiempo de cálculo de PI en sistema anfitrión, ejecutando Slackware 12	46
IV. Tiempo de cálculo de PI para un Sistema Alain en Slackware 12	47
V. Tiempo de cálculo de PI para sistema anfitrión, utilizando Windows XP SP 2	48
VI. Diferencia de rendimiento de CPU entre Sistema Alain y sistema anfitrión	49
VII. Rendimiento CPU sistema anfitrión	52
VIII. Rendimiento Sistema Alain sobre un sistema anfitrión, utilizando Slackware 12	53
IX. Rendimiento de un Sistema Alain sobre un sistema anfitrión,. utilizando Windows XP SP 2	54
X. Rendimiento de disco del sistema anfitrión, utilizando Slackware 12	56
XI. Rendimiento de disco del Sistema Alain sobre Slackware 12	57
XII. Rendimiento del Sistema Alain sobre Windows XP SP 2	58
XIII. Diferencia de rendimiento de disco del Sistema Alain y el sistema anfitrión	59
XIV. Rendimiento disco sistema anfitrión	62
XV. Rendimiento Sistema Alain sobre un sistema anfitrión, utilizando Slackware 12	64
XVI. Rendimiento de un Sistema Alain sobre un sistema anfitrión, utilizando Windows XP SP 2	66

GLOSARIO

Tópico	Descripción
Controlador	Es una pieza de software encargada de crear una abstracción de un dispositivo de hardware y ofrecer una interfaz estandarizada de este al sistema operativo de modo que este le pueda utilizar, los controladores son específicos del hardware y del sistema operativo.
Distribución Alain	Distribución de Linux basada en TTY Linux, recompilada haciendo uso del toolchain de Slackware 12.
Hipervisor	Es una capa de software que proporciona un servicio de virtualización y que corre directamente en el hardware y por debajo de un conjunto de servicios de virtualización y no sobre un sistema operativo, algunas veces se le llama implementación sobre metal desnudo.
Kernel	Es el componente central de un sistema operativo, es el responsable de administrar el conjunto de recursos de un sistema, esto es la comunicación entre los diferentes dispositivos de software y hardware. El kernel provee la capa de abstracción mas baja para interactuar con los recursos (especialmente la memoria, el procesador y los dispositivos de entrada y salida) que ofrece un sistema.
Linux	Es un sistema operativo similar a los sistemas Unix, esta formado del núcleo Linux y un conjunto de herramientas GNU, suele distribuirse en lo que se llama distribución de Linux esto es un conjunto de herramientas para administrar el sistema operativo así como un conjunto de paquetes que se pueden instalar en el sistema.
Máquina virtualizada	Contenedor de software que implementa una abstracción de un conjunto de dispositivos de hardware sobre el cual puede correr un sistema operativo.

Paravirtualización	Es un tipo de implementación de la virtualización en la cual se ofrece como máquina virtual una abstracción de hardware que requiere de un sistema operativo modificado para interactuar con el hipervisor.
Sistema anfitrión	Máquina que ofrece servicios que permite a aplicaciones ejecutarse dentro de algún tipo de contenedor que les aísla de la máquina anfitriona.
Sistema Alain	Solución formada por una máquina virtual corriendo en vmware que hace uso de la distribución Alain.
Sistema operativo invitado	Sistema operativo que correr sobre una máquina virtual y que de otro modo correría sobre una máquina física separada.

RESUMEN

La computación como todas las ciencias ha sufrido cambios debido a los avances tecnológicos, estos avances pueden proporcionar solución a problemas existentes lo que permite aplicar nuevos conceptos que representan avances y en otros casos pueden representar un problema para el modo en el que funcionan los sistemas.

En los últimos años la tecnología ha proporcionado una serie de tecnologías que han permitido crear computadores mas poderosos al mismo tiempo que ha disminuido el costo de los mismos, esto ha conducido a que las organizaciones hagan uso dedicado de los computadores a una tarea o servicio como lo puede ser un servidor de correo o base de datos, computadora de escritorio de la secretaria o un desarrollador, esto sin embargo ha tenido como efecto colateral el aumento de los costos indirectos, debido a el mantenimiento de una mayor cantidad de computadores y el hardware y software necesario para interrelacionarlo así una menor tasa de utilización de los recursos que los computadores ofrecen.

La tecnología de virtualización ofrece una solución que permite hacer uso de los avances de la capacidad y costo al tiempo que permite disminuir los costos indirectos al permitir a un computador realizar más de una tarea.

El presente trabajo explora el uso de la tecnología de virtualización para la maximización del uso de los sistemas de computadoras, así como la disminución de los costos indirectos derivados de su uso.

OBJETIVOS

GENERAL

Crear una solución basada en software libre y la tecnología de virtualización que permita:

- Minimizar el tiempo de implementación.
- Minimizar el tiempo de configuración.
- Minimizar el tamaño de la solución.
- Minimizar el costo de la solución.
- Maximización de la utilización de los recursos computacionales.

ESPECÍFICOS:

1. Creación de una distribución minimal de Linux.
2. Realización de mediciones del rendimiento de una solución de este tipo que permitan el pronostico del rendimiento de un sistema en un ambiente virtualizado.

LÍMITES

- Creación de la distribución minimal de Linux.
- Evaluacion de los distintos aspectos de performance del sistema virtual, estos son: procesador, memoria y acceso a disco.
- Se hicieron pruebas unicamente dentro del software de vmware Workstation 5.1.

INTRODUCCIÓN

El progreso en la capacidad de cálculo de las computadoras ha permitido que las máquinas con las que contamos en los escritorios tengan la capacidad de cálculo que antes tenían centros de cálculo completos.

Esta gran capacidad de cómputo es explotada por las aplicaciones para ofrecernos ambientes gráficos, mayor velocidad de cálculo, mayores capacidades de almacenamiento y por consiguiente mayor comodidad.

Sin embargo, este aumento de capacidad de cómputo se ha visto acompañado de un aumento en el consumo eléctrico, generación de calor, necesidades de refrigeración entre otros; si la capacidad de cómputo no es aprovechada del todo, este aumento se traduce en gastos directos e indirectos para el dueño del equipo.

La tecnología de virtualización ofrece la oportunidad de hacer uso de estos recursos no utilizados para ejecutar ambientes virtuales dentro de los cuales se pueden ejecutar servicios.

1 INNOVACIÓN

Las nuevas soluciones, los productos con los que interactuamos día a día experimentan un crecimiento de funcionalidades al tiempo que su costo se mantiene o disminuye lo que ha llegado a beneficiar a todas las personas.

Los nuevos teléfonos celulares sirven además de cámaras digitales, reproductores de música, grabadores de video, organizadores, clientes de correo electrónico, unidades GPS y exploradores Web todo esto en el mismo espacio que el modelo anterior de celular que poseíamos (y en algunos casos en menos espacio).

Los nuevos computadores cada vez integran más tecnologías que nos permiten comunicarnos de un modo más efectivo, las laptops ahora incluyen una cámara web, micrófono, lector de huellas digital, tarjetas de video con capacidades de aceleración de gráficos más poderosos, CPUs más poderosos, que permiten la implementación de algoritmos que a su vez permiten comunicaciones más rápidas y de más calidad.

Los nuevos coches integran una serie de tecnologías que permiten ir manejando mientras vemos en una pantalla el reporte del tiempo, así como el pronóstico del tráfico y la ruta ideal a seguir para llegar a nuestro destino en el menor tiempo posible, al tiempo que integran el equipo de sonido con el bluetooth del celular de modo que al recibir una llamada, el equipo de sonido baja el volumen y nos permite oír la llamada por los altavoces del vehículo.

Sin embargo, esta nueva funcionalidad no es siempre del todo necesaria, o puede ser el producto de motivos publicitarios más que de diseño concienzudo, de modo que nos vemos enfrentados a dispositivos tecnológicos que son cada vez más difíciles de utilizar, dificultad que se traduce en la pobre o nula utilización de las funcionalidades que presta un dispositivo, de modo que se anula la ventaja del dispositivo con respecto del anterior, por motivos del incremento en la dificultad de uso.

Ante esta situación podemos tomar dos caminos, la implementación de soluciones que no son de tipo evolutivo, esto es romper con el ciclo de aumento de la funcionalidad y volcarnos a la búsqueda de soluciones más sencilla que implementan un conjunto de funcionalidad básica llenando siempre la necesidad de un usuario, o realizar un diseño de la solución teniendo en cuenta como punto de partida el usuario del sistema para realizar un producto de diseño céntrico en el usuario.

Existen muchas compañías que han sabido aplicar este conocimiento y son conocidas por sus productos que se han convertido en el punto de referencia para otros dispositivos que eventualmente terminan copiando lo que después parece la solución obvia debido a la naturalidad que ofrece para interactuar con el dispositivo, esto es un claro indicador de un buen diseño, entre estos productos podemos mencionar:

- iPhone: Mientras que supone un incremento en la funcionalidad se aplicaron ambos principios al producir una solución de tipo disruptiva en lo que respecta a como interactuamos con el dispositivo, ya que se eliminaron todos los botones con excepción del botón de encendido, la interacción con el dispositivo se lleva a cabo, mediante la pantalla que cuenta con una tecnología táctil, esto elimina botones con rótulos y permite retroalimentar al usuario adecuadamente de las acciones que esta llevando a cabo al hacer uso de indicadores más visuales y explicativos

Figura 1. Iphone de Apple



- Ipod: Su diseño minimalista donde la usabilidad del Ipod y los productos complementarios ha sido uno de los principales puntos de preocupación fue copiado por una gran cantidad de productos rivales que cada vez se parecen más al Ipod.

Figura 2. Ipod de Apple



- Ratón de los computadores Mac: Estos tienen un único botón, esto elimina confusión con respecto a que hace cada botón y la necesidad de recordar que acción esta asociada a que botón creando una interfaz homogénea para todos los programas.

Figura 3. Mouse de una IMac



- Discos duros sólidos: Mientras que los fabricantes de discos duros tradicionales han tenido que aumentar constantemente la cache, ofrecer productos que giran a mayor velocidad y crear nuevos modos mas efectivos de guardar y recuperar la información, los discos duros sólidos que están hechos únicamente de memoria no volátil han ido ganando terreno debido a ventajas como costo uniforme de acceso a disco debido a su naturaleza, altamente resistentes a los golpes debido a su falta de partes mecanizas, eliminación de problemas de retraso de velocidad debido a problemas de desfragmentacion y menor consumo de electricidad, esto les ha valido ser usados cada vez mas en la industria de los computadores portátiles a pesar de ofrecer una menor capacidad de almacenamiento.

Figura 4. Disco duro sólido



- OLPC: Mientras que cada vez disponemos de computadores más veloces, con más memoria y más capacidad de almacenamiento existen soluciones que buscan no hacer uso de esa tendencia para proveer soluciones a un conjunto de clientes específicos. One Laptop Per Children es una iniciativa para llevar computadoras a los niños de los países subdesarrollados del mundo, las laptops están basadas en una serie de tecnologías que no son de punta, mas bien disruptivas, pero ofrecen ventajas que en conjunto permiten al diseño ofrecer ventajas únicas que le permiten desarrollarse en el ambiente para el que fue pensado y disminuir el costo, el costo proyectado de OLPC al inicio fue de US \$100, en estos momentos el costo ha aumentado a \$185 sin embargo sigue siendo una solución económica que se puede llevar a los clientes objetivos.

Figura 5. One laptop per children



Algunos detalles técnicos de OLPC son:

- a. CPU: AMD Geode GX2-500 a 366 Mhz, con un consumo de 1.0W que le permite funcionar una gran cantidad de tiempo con la batería suministrada.
- b. Memoria DRAM: 128MB DDR266.
- c. Almacenamiento secundario: Disco duro sólido de 512MB NAND flash, que le permite ser resistente a los golpes, además que es barata.
- d. Pantalla de 151.6 mm por 113.4 mm con una resolución de 1200 (H) 900 (V).
- e. Teclado: 70+ teclas
- f. Cargador de batería integrado como un mecanismo de manivela que traduce el esfuerzo físico en energía con un pequeño dinamo.
- g. Sonido, tarjeta inalámbrica, touchpad, batería de cinco celdas de 22 wats.

Sin embargo, este tipo de productos son resultado de cambios en el diseño, creación de nuevas tecnologías y el estado de la industria en ese momento.

Dentro de los factores que pueden afectar la Industria tenemos que tener en cuenta al menos:

- Si la industria se encuentra madura y los competidores están establecidos o si la industria es nueva.

Una industria madura con competidores establecidos es un campo difícil para competir, la generación de productos en este ambiente debe de ser disruptiva, de modo que se genere una nueva industria en la cual se pueda competir debido a la falta de competidores.

- El costo de la producción del producto, en industrias donde el costo de producción es alto las probabilidades de entrar a la industria son pocas, producir un producto en esta industria además enfrenta el problema de que los competidores establecidos de la industria pueden bajar el costo de su producto el tiempo que sea necesario para eliminar la competencia, de aquí que el único modo de competir en esta industria sea la generación de productos disruptivos que suplen únicamente algunas necesidades de los clientes pero tienen otras ventajas como menor costo de adquisición u operación.

1.1 Innovación de soluciones: Máquina virtual + Linux minimal = Servicio Virtual

La solución propuesta consiste en el uso de las tecnologías de virtualización, el sistema operativo Linux y el aprovechamiento de las características de cada uno de ellos para desarrollar un nuevo sistema que ofrezca una solución de tipo disruptiva con respecto a los siguientes criterios:

- Tamaño: tener el menor tamaño posible
- Servicios que ofrece: Realiza una única función.
- Recursos: diseñada para hacer el menor uso posible de recursos, incluidos RAM, Tiempo de procesador y espacio en disco duro.
- Despliegue de la solución: Implementada dentro de un contenedor virtual que se ejecuta en un equipo de cómputo que ya ofrece algún servicio.

En el capítulo 2, se realizará una introducción a la tecnología de la virtualización, sus alternativas, ventajas y desventajas.

En el capítulo 3, se realizará una introducción a las características del sistema operativo Linux, así como la personalización de este para ajustarse a los objetivos del sistema propuesto.

En el capítulo 4, se realizará un análisis del rendimiento computacional del sistema propuesto.

2 VIRTUALIZACIÓN

La virtualización es una de las tecnologías que permitirá la implementación de la solución de un modo innovador al ofrecer una característica que permite la implementación de un modo rápido y transparente, esta característica es la independencia del sistema del equipo de cómputo donde se le instale, permitiendo que el sistema de Servicio Virtual haga uso de los recursos sin uso del sistema, creando una relación donde ambos sistemas ganan y la influencia de uno sobre el otro es minimizada.

2.1 Qué es la virtualización

En computación, virtualización es un término muy amplio que se refiere a la abstracción de los recursos computacionales, una definición practica seria: “Es una técnica para esconder las características físicas de los recursos computacionales del modo en que otros sistemas, aplicaciones o usuarios finales interactúan con estos recursos. Esto incluye hacer de un solo recurso físico como un servidor, un sistema operativo, una aplicación o un recurso de almacenamiento un recurso que funcione como múltiples recursos lógicos; así también puede hacer de un conjunto de recursos físicos como servidores o recursos de almacenamiento funcionar un solo recurso lógico”.

Lo común en cualquier implementación de virtualización es el ocultamiento de detalles del recurso a virtualizar a través de la encapsulación de los

dispositivos lógicos dentro de una capa de abstracción.

El término fue usado desde 1960 y ha sido aplicado a muchos aspectos de la computación, desde sistemas completos a recursos individuales de un equipo como lo pueden ser el CPU, el almacenamiento y la RAM.

La virtualización crea una interfaz externa que oculta una implementación específica, permitiendo multiplexar para poder acceder simultáneamente un recurso, acceder recursos en múltiples localidades físicas como un solo recurso lógico o ofrecer una interfaz simplificada para el control de un sistema.

2.2 Historia de la virtualización

La virtualización fue implementada por primera vez hace más de 30 años por IBM como un modo de poder compartir lógicamente una computadora de tipo Mainframe en varias máquinas virtuales.

Estas máquinas virtuales permitían ejecutar múltiples tareas: esto es muchas aplicaciones y procesos ejecutándose al mismo tiempo, de modo que el alto costo de este tipo de máquinas se podía justificar al aprovechar mejor su capacidad.

Después durante las décadas de 1980 y 1990 se abandonó la virtualización, debido a la disminución de los costos de los servidores y las máquinas de escritorio, se produjo un cambio de arquitectura de una capa a arquitecturas de dos y n capas, lo que vino a ser el estándar para el modelo de aplicaciones distribuidas. En vez de centralizar los recursos en una servidor de tipo mainframe las organizaciones hicieron uso de máquinas de bajo costo para

implementar una gran cantidad de máquinas a modo de islas que poseían su propia capacidad de computación.

Posteriormente, la amplia adopción de Windows como sistema operativo de escritorio y la adopción progresiva de Linux en el área de servidores hicieron de la plataforma x86 el estándar de la industria.

Sin embargo, la nueva arquitectura trajo consigo nuevos retos que con el aumento del rendimiento de los equipos modernos y la disminución constante de costos ha sido únicamente exasperado, entre ellos tenemos:

- Baja utilización de la infraestructura, debido a que se suele correr una sola aplicación en un servidor, para evitar el riesgo de que vulnerabilidades en una aplicación afecten otras aplicaciones en el mismo servidor.
- Aumento del costo de mantenimiento de la infraestructura, debido a que el aumento físico de la infraestructura aumenta invariablemente el costo de mantenimiento entre otros motivos, debido a que el equipo debe permanecer operacional todo el tiempo traduciéndose esto entre otros en gastos de corriente eléctrica, enfriamiento y costos de resguardo físico (edificio, cuarto, bodega) sin que esto dependa de los niveles de utilización.
- Aumento de los costos de administración, debido a que los ambientes tienden a tornarse más complejos esto provoca que los profesionales que deben administrar estos ambientes deben poseer mayor educación y experiencia, aumentando del mismo modo los costos del personal.

- Las organizaciones gastan gran cantidad de recursos en el mantenimiento de su infraestructura desproporcionada con el nivel de uso de la misma, lo que representa más gastos.
- Alto mantenimiento de los equipos de escritorio debido a la cantidad de retos que presenta la seguridad y administración de los equipos de escritorio.

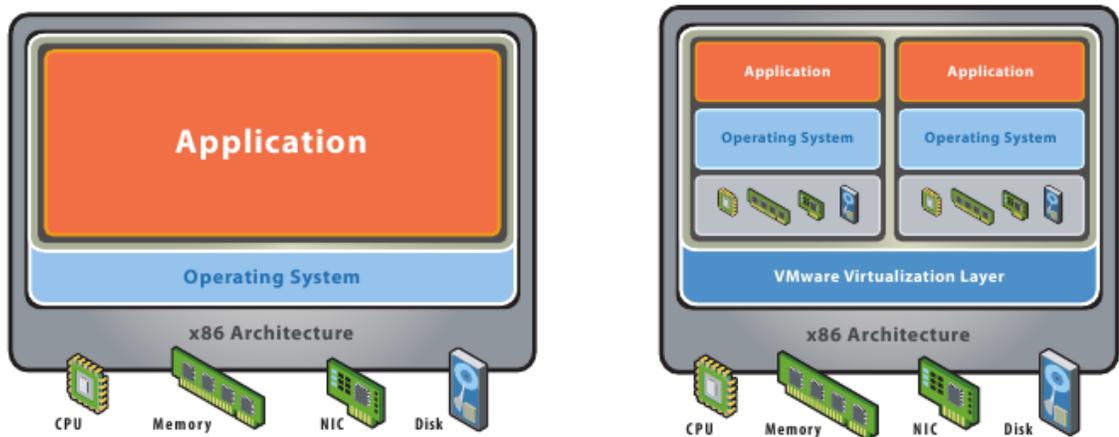
El control de un ambiente distribuido es complejo debido a que las políticas de seguridad que sea necesario implementar no deben de limitar la capacidad de trabajo de la gente.

Así mismo una gran cantidad de parches y mejoras deben de ser aplicadas continuamente a los ambientes de escritorio, para eliminar posibles fallas de seguridad.

2.3 Cómo funciona la virtualización

El modo de funcionamiento de la virtualización es proveer una interfaz que abstraer la utilización de un recurso, de modo que un servicio o aplicación pueda hacer uso de estos recursos sin estar directamente relacionada con la implementación física de éstos, es el caso del funcionamiento de la memoria virtual en la cual una parte es ubicada en almacenamiento primario y otra parte es implementada como un archivo en almacenamiento secundario, sin embargo las aplicaciones pueden hacer uso de la totalidad de la memoria sin saber acerca del funcionamiento o administración de esta a bajo nivel.

Figura 6. Arquitectura de aplicaciones en ambientes no virtualizados vs. ambientes virtualizados

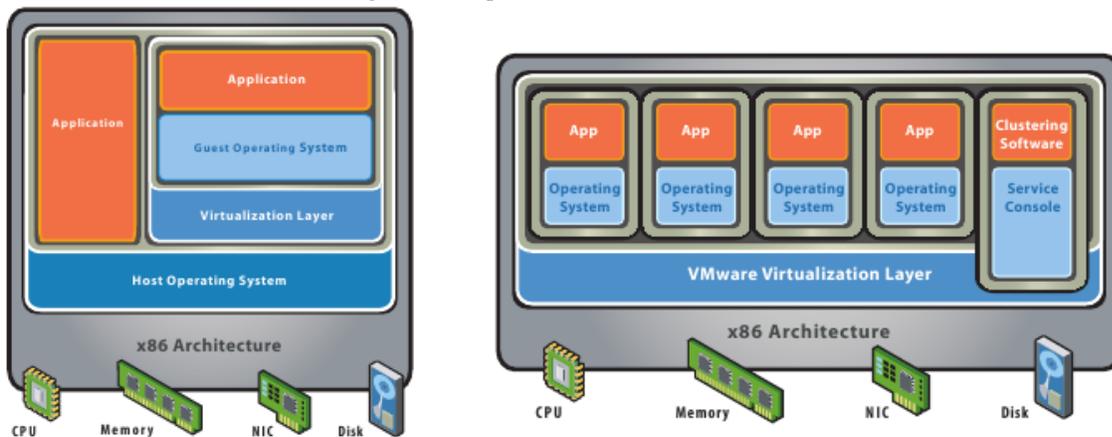


Fuente: Virtualization white paper pág. 3

Las tecnologías de virtualización nos permiten crear máquinas virtuales que funcionan del mismo modo, creando una capa de abstracción entre los recursos del equipo anfitrión y los equipos invitados, de modo que se permite la utilización de los recursos de la máquina anfitriona, por múltiples equipos invitados, esto es la multiplexación de los recursos de la máquina anfitriona para la utilización de los equipos invitados a través de la capa de abstracción.

Gracias a esta capa de abstracción cada una de las máquinas puede disponer de una serie de dispositivos virtuales, esto es dispositivos que se simulan en software y que son la interfaz que ofrece la capa de abstracción para poder acceder a los recursos de la máquina anfitriona y que para el sistema invitado tienen la forma de dispositivos de hardware como memoria, procesador, placa madre, almacenamiento en disco, tarjeta de sonido y tarjeta de red.

Figura 7. Arquitecturas de virtualización



Fuente: Virtualization white paper pág. 4

Esta capa de abstracción puede ser implementada de dos modos:

- Sobre un sistema operativo como una aplicación que corre sobre él, a este tipo de aproximación se le suele llamar “hospedado”. Este tipo de aproximación se apoya sobre el sistema operativo para la interacción con los recursos físicos del equipo.
- Con la arquitectura de hipervisor, directamente sobre la máquina, sin ningún tipo de sistema operativo; a este tipo de implementaciones suelen llamarse “sobre el metal desnudo”, en este tipo de implementación la

capa de virtualización se hace cargo de la interacción con los recursos del sistema físico, permitiendo optimizaciones en el modo de interactuar con él.

- Las arquitecturas de hipervisor pueden ser de dos tipos:
 - h. Altamente acopladas con la máquina virtual: En esta aproximación el sistema operativo que corre en la máquina virtual esta consciente de estar dentro de una máquina virtual y aprovechar este conocimiento para interactuar mas eficientemente con el hipervisor, la desventaja es que la máquina virtual no se puede implementar en un servidor físico de nuevo.
 - i. Agnóstico a la máquina virtual: En esta aproximación el sistema operativo que corre dentro de la máquina virtual no sabe que esta dentro de una máquina virtual y puede incurrir en penalizaciones de rendimiento, por este motivo, tiene la ventaja que la máquina virtual se puede implementar en un sistema físico.

2.4 Qué es una máquina virtual

Una máquina virtual es un contenedor de software que puede correr su propio sistema operativo y aplicaciones como si se tratara de una máquina física. Una máquina virtual se comporta del mismo exacto modo que lo hace una máquina física y contiene sus propios dispositivos virtuales (esto es hechos de software) como CPU, RAM, disco duro e interfaz de red (NIC).

Un sistema operativo no puede distinguir entre una máquina virtual y una máquina física, ni lo pueden hacer aplicaciones o otras computadoras en una

red. Es más la máquina virtual se comporta como si fuera una computadora real, sin embargo una máquina virtual esta compuesta de software y no contiene ningún tipo de componente de hardware. De esta cuenta una máquina virtual ofrece un número de ventajas sobre las máquinas físicas.

2.5 Ventajas de una máquina virtual

2.5.1 Compatibilidad

Del mismo modo que lo hace una máquina física, una computadora virtual sirve de anfitrión a un sistema operativo y sus aplicaciones y tienen todos los componentes que posee una máquina física (tarjeta madre, tarjeta de gráficos, tarjeta de red, etc.). Como resultado de esto, las máquinas virtuales son completamente compatibles con todos los sistemas operativos de tipo x86, sus aplicaciones y controladores de sus dispositivos, de modo que uno puede correr todo el software que se corre en una máquina física en una máquina virtual.

2.5.2 Independencia

Mientras que las máquinas virtuales pueden compartir los recursos de un solo computador físico, se mantienen independientes unas de las otras como si fueran máquinas físicas distintas.

La independencia es una importante razón por la cual los ambientes donde aplicaciones corren en ambientes virtualizados son superiores a las aplicaciones corriendo en un ambiente tradicional no virtualizado.

2.5.3 Encapsulación

Una máquina virtual es esencialmente un contenedor de software que agrupa o “encapsula” un conjunto completo de recursos de hardware, así como un sistema operativo y sus aplicaciones dentro de un paquete de software. La encapsulación hace las máquinas virtuales increíblemente portátiles y fáciles de administrar.

2.5.4 Independencia del hardware

Las máquinas virtuales son completamente independientes de la implementación física del hardware, por ejemplo uno puede configurar una máquina virtual con una serie de componentes virtuales (por ejemplo, CPU, tarjeta de red, controlador SCSI) que sean completamente diferentes de los componentes físicos que se utilizan en el hardware de la máquina física.

Esta característica sumada a la encapsulación permite que una máquina virtual pueda ser trasladada otra máquina de la misma arquitectura sin que esto implique algún tipo de configuración a los controladores, sistema operativos o a las aplicaciones.

2.6 Ventajas de la virtualización

- Optimización de la utilización de la infraestructura: La virtualización permite alcanzar niveles más altos de utilización de los recursos al permitir que los recursos estén a disposición de otros servicios que los puedan utilizar.

- Reducción del costo físico de la infraestructura: La virtualización permite reducir el número de servidores y hardware en un centro de cómputo, esto conlleva a disminución del uso de corriente eléctrica, enfriamiento y resguardo físico.
- Aumento de la disponibilidad y mejora en la continuidad del negocio: La virtualización permite eliminar los tiempos fuera de servicio no planeados, así mismo permite poder hacer copias de seguridad y migrar ambientes sin necesidad de realizar cambios al ambiente virtual.
- Aumento de la seguridad de ambientes de escritorio: La virtualización permite administrar y poner en producción ambientes de escritorio a los usuarios finales que pueden ser accedidos localmente o remotamente con o sin una conexión de red sobre casi cualquier computador de escritorio, máquina portátil o computadora de tipo tableta.

2.7 Problemas de la virtualización

La virtualización es una tecnología que ofrece muchas ventajas, sin embargo estas ventajas tienen su costo y este es más patente en aquellos aspectos donde la virtualización ofrece más ventajas, el hardware virtualizado.

El recurso a virtualizar por excelencia es el CPU la piedra ortogonal de un ordenador, este tiene que ser compartido entre todos los procesos que se ejecutan dentro de un computador y dentro de cada uno de los contenedores virtuales.

El acceso de red es otro punto a virtualizar y que es de alto impacto en ambientes corporativos debido a que representa el medio de comunicación de

la organización con los sistemas automatizados que utilizan.

2.7.1 CPU

La principal penalización asociada al CPU en ambientes virtualizados esta relacionada con la cache de este, esta es una optimización que viene dada por una pequeña memoria de rápido acceso (por lo general, esta trabaja a la velocidad del CPU) la cual almacena datos que serán de uso frecuente por el procesador al ejecutar los procesos.

El problema de la penalización de rendimiento viene dado por los contextos totalmente distintos que son el conjunto de procesos que corre dentro de cada uno de los contenedores virtuales, cambios que pueden estar implementados como saltos especializados llevados a cabo por el monitor de la máquina virtual VMM o como un conjunto de instrucciones especializadas dentro del CPU, independientemente de la aproximación tomada estos cambios de contextos no vienen a un bajo costo.

Dentro del CPU estos cambios de contexto tienen su precio más alto en la memoria de mapeo de traducciones o por sus siglas inglesas TLB (Translation Look-Aside Buffer), esta es una tabla que contiene los mapeos de las posiciones de memoria virtuales hacia sus posiciones físicas en la memoria RAM/cache para cada uno de los procesos, el hecho de cambiar entre contextos totalmente distintos representa para el CPU en el peor de los casos la perdida de todos estos registros y por consiguiente la necesidad de acceder a la memoria RAM cuya velocidad es varias veces inferior a la de la cache del CPU

para obtener esta información.

El alto costo de esta operación viene dado por que al momento de necesitar una página de memoria que no se halla, es necesario buscar dentro de los registros que están fuera de la cache del CPU e ir a buscar la memoria, esto representa de dos a cinco operaciones más de memoria para el acceso a una región de memoria que en el escenario de que la página buscada se halle dentro de la cache del CPU, sin embargo hay una segunda operación que es necesaria y esta viene dada por la actualización del registro de posiciones de memoria del VMM contra la tabla de posiciones de memoria del CPU.

Un problema asociado a los CPU no en el aspecto de rendimiento esta relacionado con el uso de instrucciones especializadas del CPU por programas ejecutados dentro de la máquina virtual, debido a que esto acopla fuertemente a la implementación de un programa dentro de la máquina virtual con características propias del CPU, anulando las ventajas de la independencia del hardware de la máquina virtual, la ejecución de una máquina virtual dentro de un anfitrión que no tiene soporte para instrucciones utilizadas por esta puede representar desde comportamiento erróneo hasta la imposibilidad de ejecución.

2.7.2 Acceso directo a memoria DMA

El uso de la memoria virtual por parte de las máquinas virtuales tiene un segundo problema relacionado con el acceso a dispositivos a través del acceso directo a memoria llevado a cabo por algunos dispositivos, tecnología mas conocida por sus siglas en ingles DMA (Direct Memory Access), esta es una optimización llevada a cabo para liberar al CPU de la tarea de tomar datos de un dispositivo hacia la memoria o de la memoria hacia el dispositivo el cual

efectivamente aumenta el rendimiento de este.

Sin embargo, la memoria dentro de las máquinas virtuales tiene una dirección física asociada la cual de ser referenciada por dispositivos que hacen uso de DMA debe de ser “pineada” esto significa que la posición de memoria física no debe de cambiar, esto en un ambiente virtual representa que no debe de cambiar dentro de la memoria de la máquina virtual y dentro de la memoria del sistema operativo anfitrión, esto complica la administración de la memoria para la máquina anfitriona al no poder relocalizar páginas de memoria asignadas a máquinas virtuales que serán referenciadas por dispositivos a través de DMA.

2.7.3 Optimizaciones de administración de memoria

Otro problema relacionado con el pineo de las posiciones de memoria tiene que ver con las optimizaciones efectuadas por el sistema operativo, para manejar grandes cantidades de memoria como una sola posición, esto disminuye la necesidad de acceso a las tablas de traducción de memoria y por consiguiente, representa un aumento de rendimiento.

Para que este aumento de rendimiento sea efectivo en ambientes virtuales es necesario que las porciones de memoria a representarse de este modo sean trabajadas como tales dentro de la máquina virtual, así como dentro del sistema operativo anfitrión, esto representa una coordinación necesaria entre ambos kernels, a través del VMM.

2.7.4 Acceso a red

Para poder proveer acceso a la red a todas las máquinas virtuales es necesario poder multiplexar el tráfico estas para poder enrutarlo a través de la interfaz o interfaces de red de la máquina anfitriona, del mismo modo es necesario demultiplexar el tráfico que es recibido a través de la interfaz de red de la máquina anfitriona para poder transmitirlo a las máquinas virtuales.

Existen varias aproximaciones para poder implementar esta arquitectura, la aproximación tomada por XEN por ejemplo, es la de definir un dominio de I/O (input output por sus siglas en inglés) que es el único encargado de comunicarse con la interfaz del sistema anfitrión al mismo tiempo que se comunica con un proveedor de interrupciones para las interfaces virtuales y provee una serie de drivers de fondo con los que se comunican las distintas interfaces de las máquinas virtuales.

La dinámica consiste en que al momento de recibir un paquete dirigido a alguna de las máquinas virtuales, el VMM se encarga de enrutar el paquete hacia el dominio de I/O este a su vez se encarga de llevar el paquete a el driver de fondo correspondiente a la máquina virtual a la cual se dirige, después se comunica con el proveedor de interrupciones para avisar que hay un nuevo paquete, en ese momento la interfaz de la máquina virtual se comunica con el driver de fondo del dominio de I/O para recibir el paquete, la dinámica es similar para el envío de un paquete.

El principal problema asociado a esta aproximación se encuentra en la latencia de los paquetes de red, así como la administración de cola que se

utilice para enlutar los paquetes de las máquinas virtuales a través del dominio de I/O para acceder a la tarjeta de red de la máquina anfitriona, corolario se encuentra el problema de la complejidad asociada a el dominio de I/O ya que debe de insular a cada una de las máquinas virtuales del acceso a la tarjeta de red de la máquina anfitriona y de entre las interfaces de red de las máquinas virtuales entre ellas.

Una solución a esto esta dado por las tarjetas de red que soportan múltiples colas, permitiendo que cada una de las tarjetas de red de las máquinas virtuales sea asignada a una de las colas definidas por la interfaz de red de la máquina anfitriona, esto simplifica el diseño al eliminar al dominio de I/O al mismo tiempo que disminuye la latencia al eliminar también al proveedor de interrupciones, de modo que la comunicación entre la interfaz de red de la máquina virtual y la interfaz de red de la máquina anfitriona es directa.

La simplificación que representa el uso de tarjetas de red que soportan multiples colas tiene un costo en seguridad, ya que el tráfico de red no es más moderado por el dominio de I/O un driver alterado puede permitir acceso de lectura a alguna de las otras colas, que puede efectivamente ver el trafico de las otras máquinas virtuales.

2.8 Comparativa de distintos productos de virtualización

Tabla I. Comparativa de productos de virtualización

Producto	Desarrollador	Arquitectura de la maquina anfitriona	Arquitectura del procesador de la maquina virtualizada	Sistema operativo anfitrión	Sistema operativo soportado por la maquina virtual	Capacidad de correr cualquier sistema operativo	Controladores para el hardware virtual	Modo de operación	Tipo de licencia-miento	Velocidad relativa a la maquina anfitriona
Linux KVM	Open source	Intel IVT / AMD-V	x86 / AMD 64	Linux	Linux / Windows	Si	No	Virtualización a nivel del kernel	GPL2	Cercana a la velocidad nativa
QEMU con modulo qemu	Open source	x86 / AMD 64	La misma que la maquina anfitriona	Linux / Windows / NetBSD	Linux / Windows	Si	No	Virtualización a nivel del kernel	GPL2	Cercana a la velocidad nativa
Virtual PC 2007	Microsoft	x86 / AMD 64	x86	Windows Vista / XP	Linux / Windows	Si	Si	Virtualización	Propietario, gratuito a partir de julio 2006	Cercana a la velocidad nativa
VMWare ESX Server 3.0	VMWare	x86 / AMD 64	x86	No necesita sistema operativo anfitrión	Linux / Windows	Si	Si	Virtualización	Propietario	Cercana a la velocidad nativa
VMWare Server	VMWare	x86 / AMD 64	x86	Linux / Windows	Linux / Windows / Solaris	Si	Si	Virtualización	Propietario, gratuito	Cercana a la velocidad nativa
VMWare workstation	VMWare	x86 / AMD 64	x86	Linux / Windows	Linux / Windows / Solaris	Si	Si	Virtualización y Paravirtualización VMI	Propietario	Cercana a la velocidad nativa

2.9 Herramienta de virtualización seleccionada: Vmware

Para la implementación del Servicio Virtual se hará uso de la tecnología de virtualización de VMware: VMware Server, las características de este producto que han motivado su selección por encima de otras implementaciones de virtualización son:

- Costo gratuito para la versión VMware Server.
- VMware es una compañía líder en el campo de virtualización.
- Amplia gama de soluciones que permiten escalar la solución horizontalmente.
- Posibilidad de escalar la solución a productos con características únicas de las tecnologías de VMware Infrastructure:

j. VMotion: VMware VMotion technology es única a los productos de VMware, esta tecnología que permite la completa virtualización de servidores, almacenamiento y comunicación de red de modo que se pueda migrar una máquina virtual de un servidor a otro.

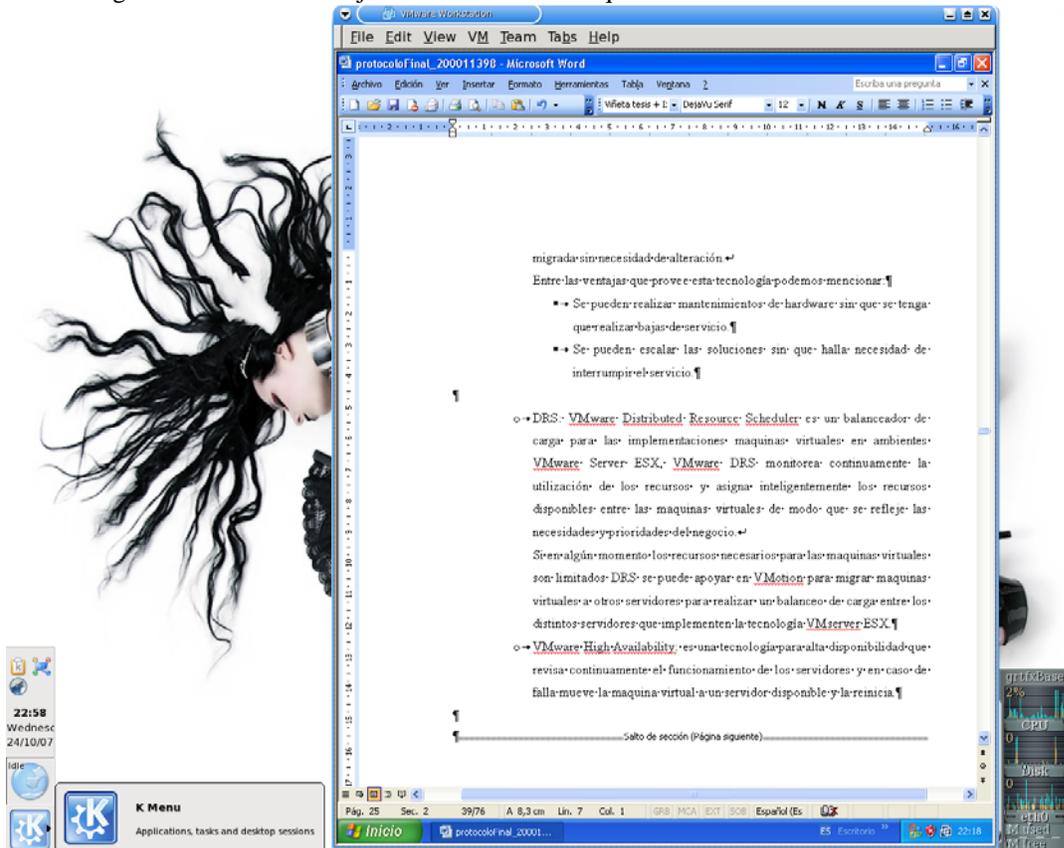
El estado de una máquina virtual es encapsulado en un conjunto de archivos en un almacenamiento compartido de modo que se puede migrar sin alterar la memoria o el procesamiento de la máquina virtual y debido a que la comunicación de red es virtualizada la máquina puede ser migrada sin necesidad de alteración.

Entre las ventajas que provee esta tecnología podemos mencionar:

- i. Se pueden realizar mantenimientos de hardware sin que se tenga que realizar bajas de servicio.

- ii. Se pueden escalar las soluciones sin que haya necesidad de interrumpir el servicio.

Figura 8. Windows XP ejecutándose en una máquina virtual en Slackware Linux 12



k. DRS: VMware Distributed Resource Scheduler es un balanceador de carga para las implementaciones máquinas virtuales en ambientes VMware Server ESX, VMware DRS monitorea continuamente la utilización de los recursos y asigna inteligentemente los recursos disponibles entre las máquinas virtuales de modo que se refleje las necesidades y prioridades del negocio.

Si en algún momento los recursos necesarios para las máquinas virtuales son limitados DRS se puede apoyar en VMotion para migrar máquinas virtuales a otros servidores para realizar un balanceo de carga entre los distintos servidores que implementen la tecnología VMware ESX.

l. VMware High Availability: es una tecnología para alta disponibilidad que revisa continuamente el funcionamiento de los servidores y en caso de falla mueve la máquina virtual a un servidor disponible y la reinicia.

3 LINUX

En el capítulo uno se exploró el enfoque de la innovación en varios productos y cómo esto les ha permitido convertirse en productos exitosos. En el capítulo dos se exploró el uso de la tecnología de la virtualización, utilizando sus características inherentes para generar un producto innovador con un enfoque distinto.

Sin embargo, la tecnología de la virtualización es únicamente el contenedor del sistema, dentro del contenedor es necesario ubicar un sistema operativo robusto y flexible, que permita implementar un sistema liviano, veloz y seguro que sirva de soporte a los servicios que proporciona el Servicio Virtual.

Este sistema operativo es Linux, un producto maduro y flexible que permite ajustarlo a las necesidades de aquello que soporta. Este capítulo explora el sistema operativo Linux, así como la personalización aplicada para ajustarse a los objetivos del sistema.

3.1 Qué es Linux?

Linux es un sistema operativo parecido a Unix. Linux es uno de los más prominentes ejemplos de software libre y desarrollo de tipo abierto, el código del que está formado, puede ser libremente modificado, usado y distribuido por cualquier persona.

Linux esta formado por el kernel Linux desarrollado bajo la guía de Linus Torvalds, un conjunto de documentación, un manejador de ventanas, un ambiente de escritorio, un conjunto de librerías y herramientas de software GNU desarrollados por la FSF, el conjunto del kernel de Linux y Software del proyecto GNU es llamado distribución de Linux, esta puede contener un kernel modificado, software GNU así como software propietario.

La primera versión del kernel de Linux fue liberada al público el 17 de septiembre de 1991, para la arquitectura Intel x86 PC. Posteriormente el kernel se empezó a acompañar de utilidades y librerías del proyecto GNU para crear un sistema operativo que pudiera ser usable lo que después condujo al uso del término alternativo GNU/Linux.

Debido a su estabilidad, escalabilidad, portabilidad de arquitectura y alto rendimiento Linux ha sido usado en una gran cantidad de sistemas que van desde celulares, routers, supercomputadores, máquinas de escritorio, handheld hasta servidores, donde ha sido su uso predominante, esto le ha valido el soporte de una gran cantidad de corporaciones como lo son:

- IBM
- Sun Microsystems
- Dell
- Hewlett-Packard
- Novell

3.2 Qué es una distribución de Linux?

Una distribución de Linux, también llamada distribución o distro es un sistema operativo GNU/Linux que contiene el kernel Linux, un conjunto de herramientas GNU y posiblemente software de otro tipo.

Las distribuciones de Linux pueden tomar una gran cantidad de formas dependiendo del mercado objetivo, estas pueden ir desde distribuciones que buscan ofrecer funcionalidad para un servidor o para una máquina de escritorio y se distribuyen en conjuntos de cds o dvds hasta distribuciones minimales que corren desde un disquete y sirven para llevar a cabo tareas de mantenimiento en un sistema.

Existen otros aspectos fuera de la funcionalidad que ofrece una

distribución que las diferencia entre si, entre ellas:

- Tipo de paquetes para instalar software en la distribución (rpm, deb, source, etc.): una distribución esta formada de paquetes que contienen software que tiene alguna finalidad, por ejemplo un paquete de software puede contener el conjunto de librerías que permiten a un programa poder interactuar con el manejo de gráficas en formato jpg.

Un paquete por lo general contiene software en formato binario, esto es compilado para la ejecución en lenguaje de máquina, así también puede contener información acerca de acciones previas a la instalación, acciones posteriores a la instalación, meta información acerca del paquete como lo puede ser dependencias que es necesario suplir previo a la instalación del paquete.

- Manejador de paquetes para instalar los paquetes en la distribución: este es el software responsable de llevar a cabo las actividades que permitan la correcta instalación de un paquete como lo puede ser ejecutar tareas previas necesarias para la instalación de un paquete como lo son establecer permisos para ejecución, resolución de dependencias, esto es instalar un conjunto de paquetes necesarios para el correcto funcionamiento del paquete que se desea instalar, crear un conjunto de reglas para el funcionamiento inicial del paquete que se va a instalar, por lo general esto se lleva a cabo mediante la creación de un conjunto de archivos de configuración; tareas posteriores a la instalación que pueden incluir creación de archivos, actualización de permisos, creación de archivos de configuración, etc.

- Medio de instalación (por ejemplo dvd, cd, disquete, cd de forma limitada como lo es un cd de 230 mb o formato de tarjeta de presentación de 50 mb, memorias de conexión usb, etc): esto depende de la necesidad del usuario objetivo, por lo general distribuciones de Linux de objetivo general suelen venir acompañadas de una gran cantidad de software por lo que se suelen distribuir en uno o varios dvds o varios cds, algunos buscan instalar la funcionalidad básica de un usuario de escritorio y pueden venir en un cd, otras especializadas en el mantenimiento pueden venir en un cd de 230 mb o un disquete.
- Escritorio por defecto (por ejemplo Gnome, KDE, fluxbox, etc).
- Arquitectura para la que está compilado el kernel, las librerías y las herramientas de software (por ejemplo, Intel x86 PC, AMD64, SPARC, etc.).
- Desarrollado con respaldo de una corporación (Ubuntu-canonical, SUSE-Novell) o no (Gentoo, Slackware).
- Idioma por defecto: muchas distribuciones de Linux son versiones regionalizadas de otras distribuciones, esto es incluye un conjunto de software que viene de forma predefinida en un idioma, pensemos en el diccionario de corrección, idioma del manejador de escritorios, etc..

3.3 TTYLinux

TTY Linux es una distribución minimal de Linux que tiene los objetivos de:

- Correr como mínimo sobre un procesador Intel 486dx
- Utilizar 5 MB de RAM al correr desde disco.
- Utilizar 4 MB para la instalación en el disco duro.
- Utilizar 10 MB de RAM al correr desde el cd.
- Poder utilizar un MODEM o conexión de Ethernet para conectarse a Internet.

Objetivos que lleva a cabo haciendo uso de un conjunto de herramientas bastante livianas como lo son busybox para las aplicaciones básicas de usuario para interactuar con el sistema operativo, Interprete de comandos BASH, kernel 2.6.16 sin soporte para módulos con lo que se gana bastante espacio además de que se vuelve bastante liviana la distribución, además de ofrecer utilidades como:

- Interfaz de usuario de tipo texto, no contiene una interfaz grafica de usuario.
- Implementación de un Firewall haciendo uso del módulo de kernel iptables.
- Navegador tipo texto.
- Cliente y servidor de conexiones seguras SSH.
- Editor de texto Vi.

Sin embargo, TTY Linux en su forma vanilla tiene muchos inconvenientes, para usarse como una solución de tipo general, entre los que he hallado:

- Falta de soporte para agregar paquetes de un modo sencillo, el manejador de paquetes provisto en una simple aplicación que descomprime archivos y los ubica en el sistema raíz.
- Limitados paquetes para utilizar con la distribución (al momento de revisar la pagina únicamente se ofrecían dos paquetes):
 - m. Un cliente de IRC tipo texto.
 - n. Una implementación del servidor tinyhttp.
- Cadena de compilación basada en Debian Sarge.

Siendo los problemas anteriores un obstáculo debido a que la falta de un manejador de paquetes robusto así como paquetes para instalar limita la capacidad de uso de la distribución, obligando a la compilación de paquetes con el toolchain de Debian Sarge.

3.4 Sistema Alain

El Sistema Alain es un nuevo sistema operativo creado a partir del sistema operativo TTY Linux que busca corregir los errores que impiden la utilización de TTY Linux y aprovechar todas las ventajas que tiene para ofrecer.

El principal problema para la utilización de TTY Linux en su forma vanilla son los pocos paquetes que hay para la distribución y el hecho de que se necesita el mismo toolchain que se usa en Debian Sarge para la compilación de los paquetes a agregar.

El Sistema Alain un nuevo sistema que da solución a los problemas de TTY Linux y aprovecha sus ventajas, esto se lleva a cabo al recompilar la distribución TTY Linux con el toolchain de Slackware, esto resuelve los tres problemas iniciales que presenta TTY Linux en su forma vanilla, haciendo uso de este toolchain para recompilar la distribución se puede hacer uso de los paquetes de Slackware para instalar en la el Sistema Alain.

Entre otros, los motivos para la selección de Slackware fueron:

- Utilización de paquetes vanilla: esto es los paquetes no han sido modificados, sino que han sido usados los fuentes suministrados por los proveedores de la aplicación.
- Compilada para arquitectura Intel x86 con un mínimo de procesador 486, con esto se pierde capacidad para la ejecución de aplicaciones que sean instaladas desde Slackware para ejecutarse en procesadores más viejos.
- Slackware es conocida por ser una de las distribuciones más rápidas y estables.
- Amplia selección de paquetes que va desde herramientas de consola, pasando por paquetes de escritorio como KDE hasta aplicaciones de servidor como bases de datos y servidores de correo.

- Utilización de un formato para los paquetes que consiste en un conjunto de archivos empaquetados en formato tar.gz, lo cual es fácilmente instalable en TTY LINUX, mediante el uso de la aplicación tar y gzip.

3.5 Modificación y compilación de Sistema Alain

El primer paso para la creación del Sistema Alain es la recopilación de la distribución, esto se lleva a cabo al bajar del sitio de TTY Linux los paquetes necesarios para crear la distribución, los paquetes necesarios son los siguientes:

Tabla II. Paquetes del Sistema Alain

Nombre del paquete	Version	Tamaño	Función
Basesystem	6.1-1	50Kb	Sistema base y conjunto de scripts para crear la distribución
Bash	3.2.17-1	1.8Mb	Interprete de comandos
Busybox	1.6.1-1	1.1Mb	Conjunto de herramientas para interactuar con el sistema operativo en línea de comandos
Dropbear	0.49-1	1.3Mb	Cliente SSH
e2fsprogs	1.40-1	3.1Mb	Conjunto de herramientas para interactuar con el sistema de archivos
Glibc	2.5-1	14Mb	Conjunto de librerías compartidas GNU
Iptables	1.3.8-1	160Kb	Conjunto de herramientas para interactuar con la pila TCP/IP del kernel de Linux
isdn4k-utils	3.2p1-4	5.1Mb	Conjunto de herramientas para establecer una comunicación sobre un medio ISDN
LILO	22.8-1	360Kb	Linux Loader es el cargador de facto de los sistemas Linux

Ncurses	5.6-2	1.6Mb	Conjunto de librerías para dibujo sobre la pantalla
PPP	2.4.4-3	548Kb	Conjunto de herramientas para establecer una conexión de tipo punto a punto
retawq	0.2.6c-3	264Kb	Navegador web tipo texto
Linux kernel	2.6.16.53	40Mb	Kernel de Linux

Con la excepción del kernel Linux todos los paquetes se pueden descargar desde el sitio de TTY Linux, para el kernel es necesario ir a el sitio del kernel de Linux.

Después de esto es necesario acomodar el sistema de generación de la distribución, así como ubicar los fuentes dentro de sus lugares respectivos, los pasos a realizar son:

- Desempaquetar el archivo ttylinux-src, este contiene los archivos necesarios para la compilación de la distribución.
- Ubicar los fuentes de los archivos que deseamos en formato .tar.bz dentro de la carpeta linuxtty/packages/src/
- Ejecutar el archivo linuxtty/makedist.sh
- Al completar la ejecución de makedist.sh se generará un archivo que contiene la distribución sin el kernel, el nombre del archivo es ttylinux-i486-6.1.tar.gz.
- Al desempaquetarlo obtendremos el directorio llamado ttylinux, dentro de este directorio se encuentra un stript llamado mkTTYiso, este es el que crea el archivo live cd, sin embargo para crear el live cd haciendo uso de este script necesitamos dos cosas más que son crear el kernel de la distribución y el cargador para el live cd, el cargador se llama isolinux.
- La compilación del kernel se llevó a cabo haciendo uso del script de

configuración del kernel que suministra la persona que mantiene la distribución, para ello se utilizó la última revisión del kernel 2.6.16 al momento que es la 2.6.16.53.

- Para producir un kernel es necesario descomprimir los fuentes del kernel y copiar dentro del primer nivel del directorio el archivo config del kernel suministrado por la persona que mantiene ttylinux y copiarlo con el nombre de .config.

Después de esto si deseáramos cambiar algo dentro del kernel se puede ejecutar `make xconfig` o `make menuconfig` para configurar el kernel a gusto.

Al terminar la configuración es necesario compilar el kernel, para realizarlo es necesario ejecutar el comando `make bzImage`.

Al terminar se obtendrá el kernel dentro del directorio `arch/i386/boot` este tendrá el nombre de `bzImage`.

- Para la generación del live cd que nos permite instalar el sistema operativo hay necesidad de descargar el paquete SYSLINUX, este paquete provee una serie de programas que permiten iniciar el sistema operativo Linux de un conjunto de medios distintos a un disco duro, por ejemplo se provee PXELINUX que sirve para iniciar el sistema operativo Linux desde el sistema PXE de red.

Para que la distribución pueda arrancar desde el cd es necesario el archivo `isolinux.bin` que es generado entre otros al compilar los fuentes de SYSLINUX, para esto es necesario descomprimir el paquete, entrar al directorio que contiene los fuentes, preparar el ambiente de compilación con el comando `make`, compilar los fuentes con el comando `make`.

Al terminar el proceso en el directorio se ubicará el archivo `isolinux.bin`, este tiene que ser copiado a la carpeta `ttyLinux` descomprimida previamente, al terminar y estar aquí ubicado el kernel es posible generar

la distribución ejecutando el script que se incluye en este directorio llamado mkttyiso, para ejecutarlo hay que escribir en una consola ./mkttyiso.

Al terminar el proceso obtendremos un archivo llamado bootcd.iso, con este archivo podemos crear un cd de arranque de ttyLinux con software para quemar cds como nero o k3b en Linux.

4 RENDIMIENTO DE SISTEMA ALAIN EN UN AMBIENTE VIRTUAL

El conjunto de tecnologías que se utilizan para la creación de la Servicio Virtual poseen características únicas que necesitan ser probadas en su conjunto en ambientes de alto estrés para asegurar que se pueda proveer una solución que sea liviana y rápida al tiempo que cumpla con los objetivos propuestos.

Los principales aspectos a medir están relacionados con el rendimiento de los recursos de los sistemas virtualizados con respecto a los recursos del sistema anfitrión, esto es los recursos que comparten las máquinas virtuales, entre otros el CPU, la memoria, la salida de red y el acceso a disco, de modo que se pueda generar una métrica de la penalización que impone la virtualización al acceso de los recursos físicos y se puedan realizar estimaciones de la capacidad necesaria para un conjunto de máquinas virtuales.

4.1 Equipo de prueba: máquina de escritorio

Las características del sistema de pruebas son:

- Ordenador de escritorio tipo clon:
 - o. Procesador AMD Athlon 64 de 2.0Ghz con 512KB de cache L2.
 - p. Memoria RAM de 1.5GB de tipo DDR a 400Mhz.
 - q. Disco duro de 80GB a 7200 revoluciones con 8MB de cache.
- Sistema operativo Slackware 12:
 - r. Compilado para arquitectura Intel 486.
 - s. Versión de kernel 2.6.21.5.
 - t. Soporte para SMP.
 - u. GCC 4.2.1
- Ambiente de virtualización VMware Workstation:
 - v. Tamaño del disco duro virtual:
 - i. 102 MB no preasignados para el sistema operativo.
 - ii. 3 GB preasignados para el disco duro de pruebas
 - w. Área de intercambio de memoria en disco virtual: ninguna.
 - x. Memoria virtual asignada: 24 MB.

4.1.1 Análisis del rendimiento de uso de CPU

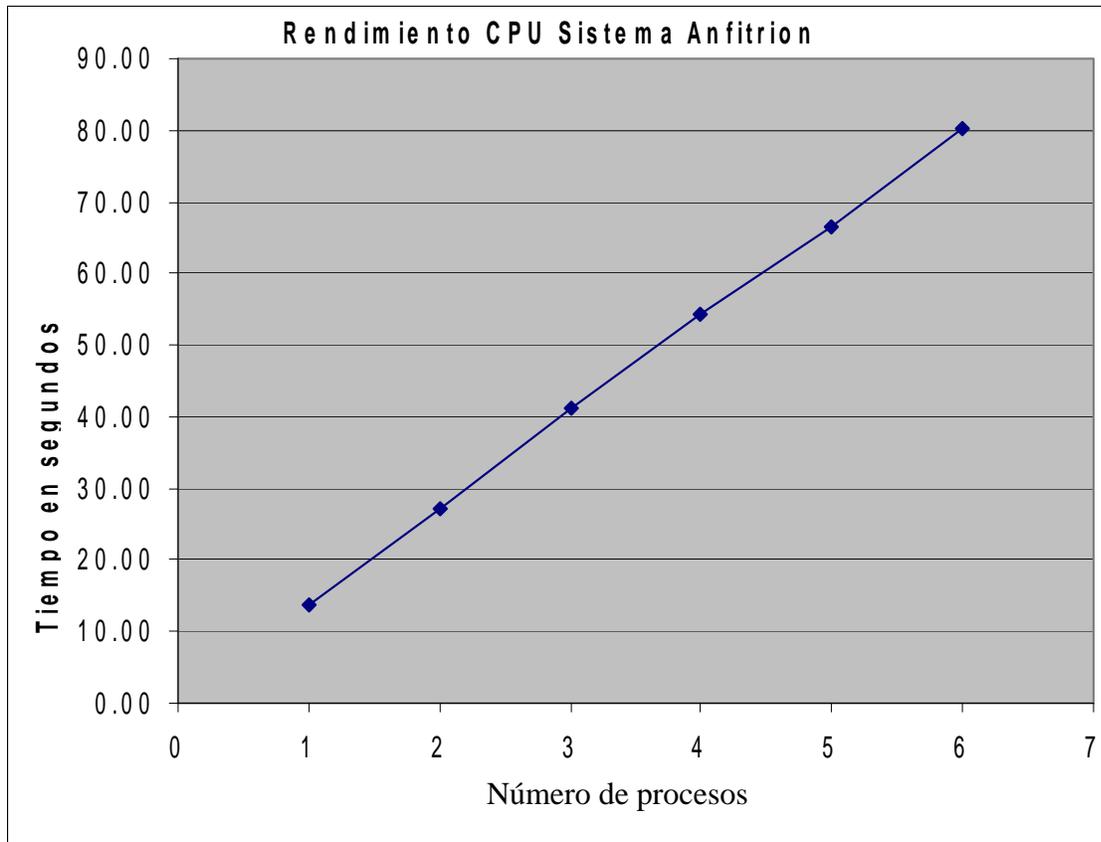
Para medir la penalización que sufre el Sistema Alain haciendo uso intensivo del CPU en un ambiente virtual se ejecutara una aplicación que hace el calculo de los primeros cincuenta mil dígitos de PI en el ambiente nativo y en el ambiente virtual, en una y en varias máquinas, lo que se busca es ver la penalización de la capa de virtualización, así como la penalización en la que incurre la infraestructura virtual para atender varias máquinas virtuales.

4.1.1.1 Resultados sistema anfitrión utilizando Slackware 12

Tabla III. Tiempo de cálculo de PI en sistema anfitrión, ejecutando Slackware 12

# procesos	tiempo promedio en segundos	Tiempo en seg para N procesos					
		1	2	3	4	5	6
1	13.71	13.82	27.05	41.19	54.75	67.09	80.50
2	27.16	13.34	26.76	41.10	54.31	66.48	80.06
3	41.06	13.71	27.59	40.88	54.03	66.22	79.95
4	54.30	13.31	27.40	40.90	54.66	66.80	80.56
5	66.58	14.09	26.96	41.10	54.25	66.49	80.22
6	80.32	13.96	27.20	41.19	53.81	66.38	80.61

Figura 9. Tiempo de cálculo de PI para el sistema anfitrión, utilizando Slackware 12

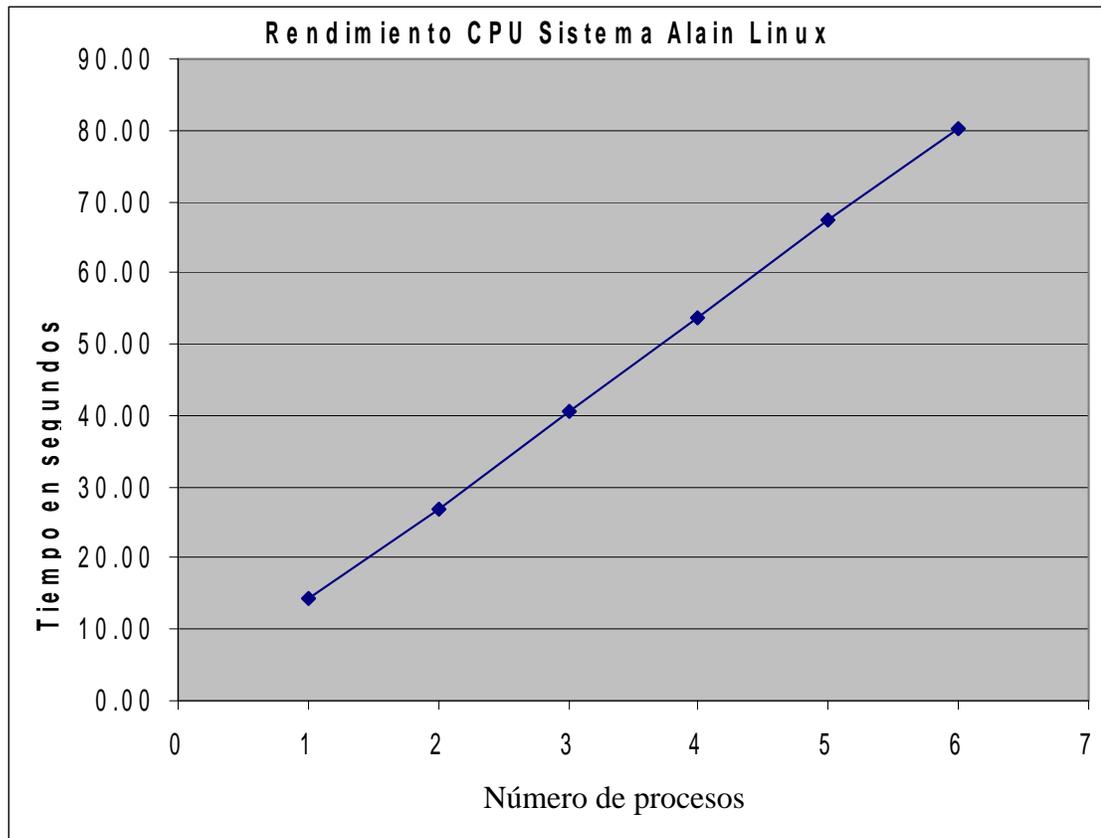


4.1.1.2 Un Sistema Alain corriendo en un sistema anfitrión, utilizando Slackware 12

Tabla IV. Tiempo de cálculo de PI para un Sistema Alain en Slackware 12

# procesos	tiempo promedio en segundos	Tiempo en seg para N procesos					
		1	2	3	4	5	6
1	14.38	14.53	27.25	40.64	53.96	67.70	80.43
2	26.94	14.06	27.00	40.66	53.44	67.60	80.60
3	40.45	13.94	26.81	40.40	53.48	66.87	80.11
4	53.66	14.64	26.62	40.60	53.59	67.02	80.28
5	67.35	14.33	26.84	40.59	53.64	67.47	80.03
6	80.34	14.78	27.14	39.78	53.89	67.45	80.61

Figura 10. Tiempo de cálculo de PI para Sistema Alain, corriendo en un sistema anfitrión utilizando Slackware 12

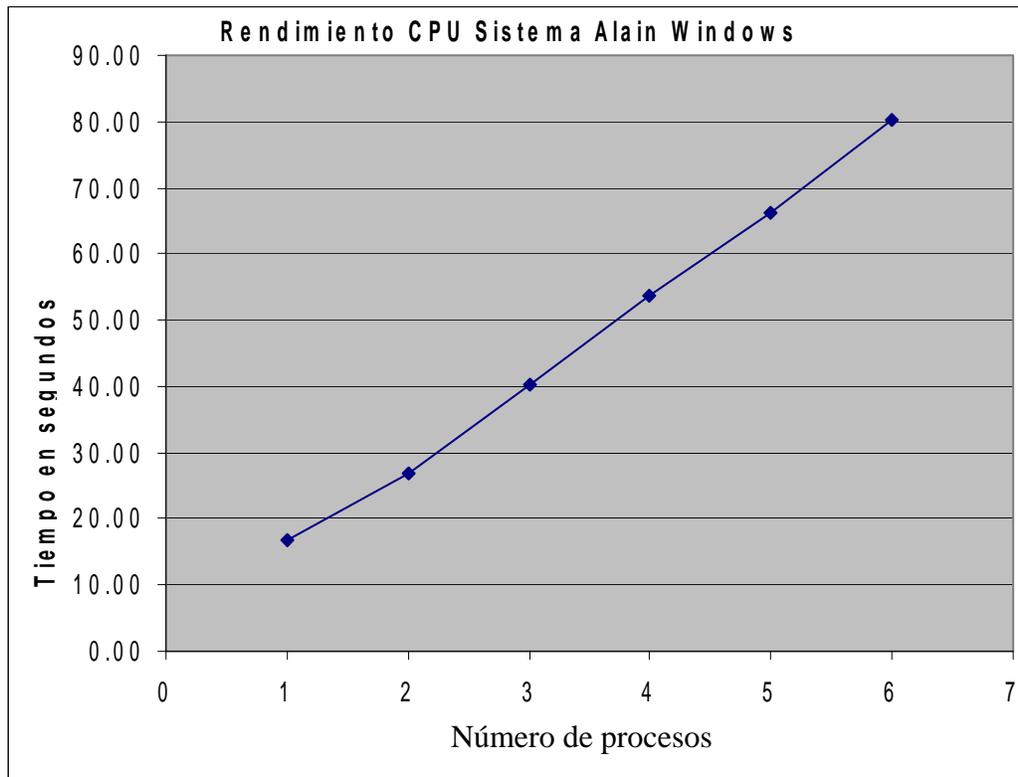


4.1.1.3 Un Sistema Alain corriendo en un sistema anfitrión, utilizando Windows XP SP2

Tabla V. Tiempo de cálculo de PI para sistema anfitrión, utilizando Windows XP SP 2

# procesos	tiempo promedio en segundos	Tiempo en seg para N procesos					
		1	2	3	4	5	6
1	16.71	16.38	26.70	39.89	53.91	65.73	80.33
2	26.88	16.82	27.07	40.55	53.86	66.40	80.30
3	40.17	16.62	27.01	40.30	53.78	66.30	80.06
4	53.62	16.69	26.66	39.76	53.30	65.50	79.73
5	66.11	16.64	26.79	39.99	53.49	66.48	80.03
6	80.14	17.12	27.06	40.51	53.39	66.23	80.41

Figura 11. Tiempo de cálculo de PI para Sistema Alain corriendo en un sistema anfitrión, utilizando Slackware 12



4.1.1.4 Diferencia de rendimiento para aplicaciones intensivas de CPU entre el Sistema Alain y el sistema anfitrión.

Tabla VI. Diferencia de rendimiento de CPU entre Sistema Alain y sistema anfitrión

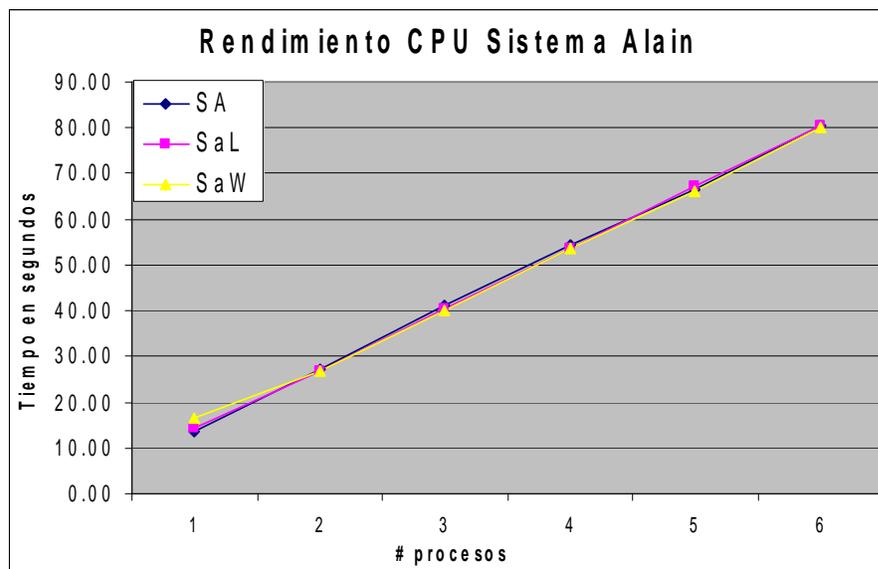
Tiempo en segundos para distintas configuraciones			
# procesos	SA	SaL	SaW
1	13.71	14.38	16.71
2	27.16	26.94	26.88
3	41.06	40.45	40.17
4	54.30	53.66	53.62
5	66.58	67.35	66.11
6	80.32	80.34	80.14

SA= Sistema anfitrión corriendo Slackware Linux 12

SaL= Sistema Alain en Sistema Anfitrión Slackware Linux 12

SaW= Sistema Alain en Sistema Anfitrión Windows XP

Figura 12. Diferencia de rendimiento para aplicaciones intensivas de CPU entre el Sistema Alain y el sistema anfitrión.



4.1.2 Predicción del rendimiento de CPU para los distintos sistemas

Con los datos adquiridos a través de las mediciones de tiempo del calculo de dígitos de PI en los distintos sistemas (anfitrión y Alain) se ha calculado haciendo uso de regresión lineal rectas que describen el rendimiento de los sistemas y lo que es mas importante la diferencia de rendimiento que supone la virtualización para el Sistema Alain.

La penalización a la que se ve expuesta el Sistema Alain sobre un sistema anfitrión utilizando Slackware Linux 12 viene dada por:

Ecuación 4-1 Fórmula para el porcentaje de penalización CPU para Sistema Alain sobre un sistema anfitrión utilizando Slackware Linux 12

$$y = \frac{0.7626667 + 13.264x}{0.733333 + 13.273x}$$

Donde Y es el porcentaje de penalización que sufrirá el sistema y
X es el número de procesos ejecutándose simultáneamente

La penalización a la que se ve expuesta el Sistema Alain sobre un sistema anfitrión utilizando Windows XP SP2 viene dada por:

Ecuación 4-2 Fórmula para el porcentaje de penalización CPU para Sistema Alain sobre un sistema anfitrión utilizando Windows XP SP 2

$$y = \frac{2.442667 + 12.80828571x}{0.733333 + 13.273x}$$

Donde Y es el porcentaje de penalización que sufrirá el sistema y
X es el número de procesos ejecutándose simultáneamente

4.1.2.1 Sistema anfitrión

Tabla VII. Rendimiento CPU sistema anfitrión

a	b	x	Medido	y lineal	%diff
0.733	13.273	1	13.71	14.01	2%
		2	27.16	27.28	0%
		3	41.06	40.55	1%
		4	54.30	53.82	1%
		5	66.58	67.10	1%
		6	80.32	80.37	0%
		7		93.64	
		8		106.92	
		9		120.19	
		10	126.64	133.46	5%

Correlacion modelo lineal	0.999861
Covarianza modelo lineal	46.455

El tiempo que le lleva a **X** procesos en paralelo dentro de un solo sistema Anfitrión completar la tarea de calcular 50000 dígitos de PI esta dado por la fórmula

Ecuación 4-3 Fórmula de tiempo para cálculo de PI en función del número de procesos en paralelo para un sistema anfitrión

$$y = 0.733333 + 13.273x$$

Donde Y es el tiempo en segundos que tomara completar la tarea y
X es el número de procesos ejecutándose simultáneamente

Donde **y** es el tiempo que le tomara al Sistema Alain llevar a cabo los **X** procesos paralelamente.

4.1.2.2 Sistema Alain sobre un sistema anfitrión utilizando Slackware 12

Tabla VIII. Rendimiento Sistema Alain sobre un sistema anfitrión, utilizando Slackware 12

a	b	X	Medido	y lineal	%diff	% overhead
0.762667	13.264	1	14.38	14.03	3%	0%
		2	26.94	27.29	1%	0%
		3	40.45	40.55	0%	0%
		4	53.66	53.82	0%	0%
		5	67.35	67.08	0%	0%
		6	80.34	80.35	0%	0%
		7		93.61		0%
		8		106.87		0%
		9		120.14		0%
		10	128.45	133.40	4%	0%

Correlación modelo lineal	0.999942
Covarianza modelo lineal	46.4240

El tiempo que le lleva a **X** procesos en paralelo dentro de un solo Sistema Alain sobre un sistema anfitrión utilizando Slackware 12 completar la tarea de calcular 50000 dígitos de PI esta dado por la fórmula

Ecuación 4-4 Fórmula de tiempo para calculo de PI en función del numero de procesos en paralelo para un Sistema Alain sobre un sistema anfitrión utilizando Slackware 12

$$y = 0.7626667 + 13.264x$$

Donde Y es el tiempo en segundos que tomara completar la tarea y
X es el número de procesos ejecutándose simultáneamente

4.1.2.3 Sistema Alain sobre un sistema anfitrión utilizando Windows XP SP 2

Tabla IX. Rendimiento de un Sistema Alain sobre un sistema anfitrión, utilizando Windows XP SP 2

a	b	x	Medido	y lineal	%diff	% overhead
2.442667	12.80828571	1	16.71	15.25	10%	9%
		2	26.88	28.06	4%	-3%
		3	40.17	40.87	2%	-1%
		4	53.62	53.68	0%	0%
		5	66.11	66.48	1%	1%
		6	80.14	79.29	1%	1%
		7		92.10		2%
		8		104.91		2%
		9		117.72		2%
		10	126.89	130.53	3%	2%

Correlacion modelo lineal	0.96765
Covarianza modelo lineal	1.92250

El tiempo que le lleva a **X** procesos en paralelo dentro de un solo Sistema Alain sobre un sistema anfitrión utilizando Windows XP SP 2 completar la tarea de calcular 50000 dígitos de PI esta dado por la fórmula

Ecuación 4-5 Fórmula de tiempo para calculo de PI en función del número de procesos en paralelo para un Sistema Alain sobre un sistema anfitrión, utilizando Windows XP SP 2

$$y = 2.442667 + 12.80828571x$$

Donde Y es el tiempo en segundos que tomara completar la tarea y
X es el número de procesos ejecutándose simultáneamente

4.1.3 Análisis del rendimiento de almacenamiento en disco

Para las pruebas del rendimiento del acceso a disco en un ambiente virtualizado realizare la descompresión de los archivos fuentes del kernel de Linux que se utilizaron para la generación del kernel que se usa en el Sistema Alain, para ello primero comprimiré los fuentes con los binarios generados en un archivo tar.gz y después los descomprimiré en el sistema anfitrión y en el Sistema Alain para medir la penalización que le impone la virtualización.

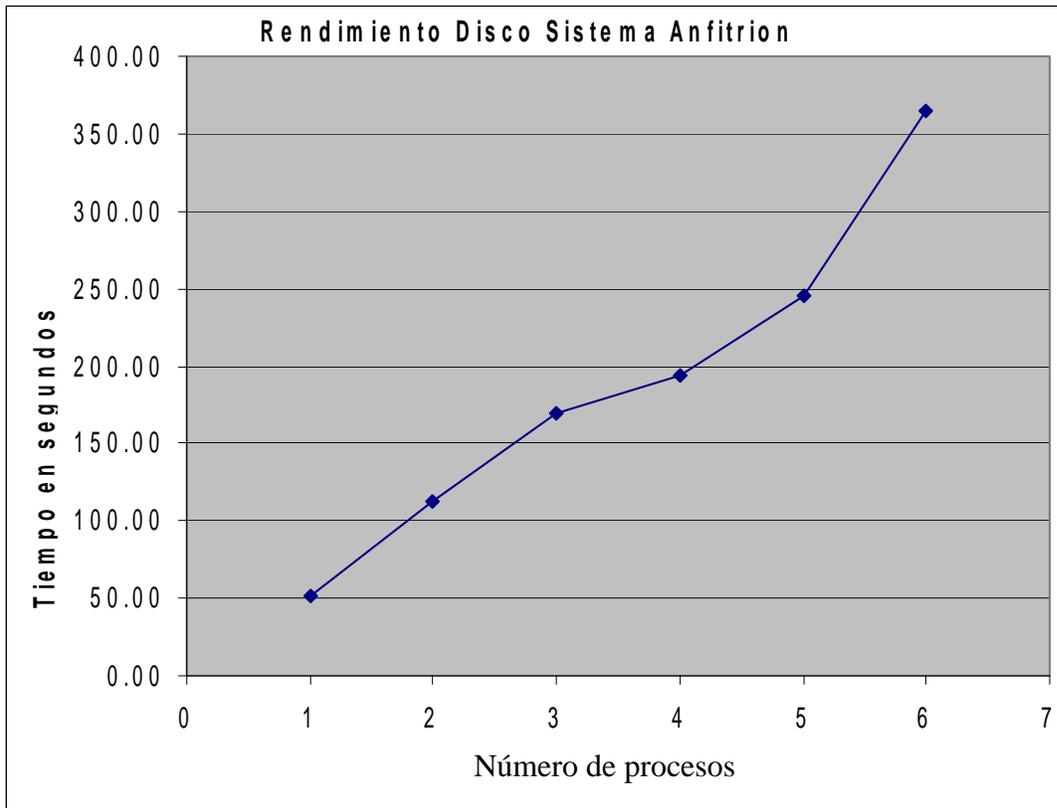
Las mediciones se realizaran haciendo uso de discos duros virtuales para los cuales previamente se ha apartado el espacio en disco.

4.1.3.1 Resultados sistema anfitrión, utilizando Slackware 12

Tabla X. Rendimiento de disco del sistema anfitrión, utilizando Slackware 12

# procesos	tiempo promedio en segundos	Tiempo en seg para N procesos					
		1	2	3	4	5	6
1	51.09	51.26	112.6	168.6	194.3	246.1	365.1
2	112.89	51.36	113.5	168.7	193.5	245.8	365.2
3	168.91	50.75	112.6	169.5	194.1	245.6	365
4	193.98	50.64	113.1	168.7	194.1	245.8	365.1
5	245.79	51.25	112.9	169.4	194.3	245.6	364.6
6	365.01	51.29	112.7	168.6	193.6	245.9	365.1

Figura 13 Tiempo de descompresión para el sistema anfitrión, utilizando Slackware 12

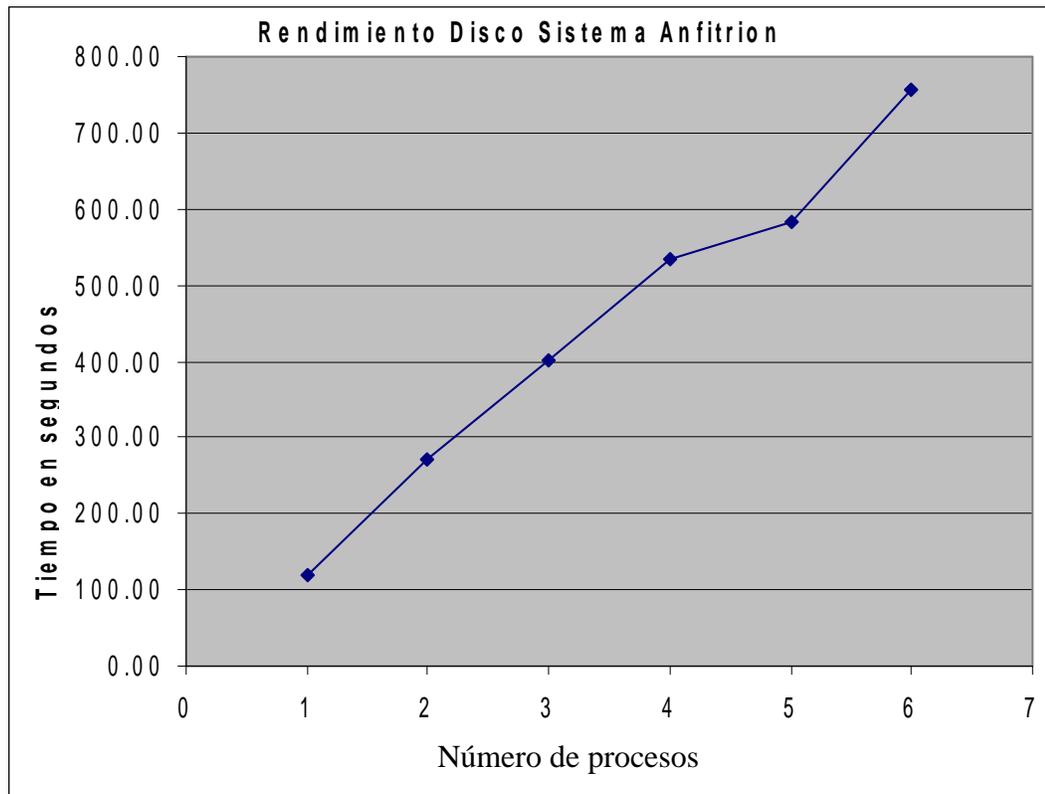


4.1.3.2 Resultados Sistema Alain sobre un sistema anfitrión, utilizando Slackware 12

Tabla XI. Rendimiento de disco del Sistema Alain sobre Slackware 12

# procesos	tiempo promedio en segundos	Tiempo en seg para N procesos					
		1	2	3	4	5	6
1	118.64	118.48	271.93	402.28	528.05	575.44	769.90
2	271.80	119.00	272.31	402.88	535.09	582.76	745.98
3	402.07	118.47	272.24	399.40	541.55	591.14	756.63
4	535.21	119.58	271.72	401.58	533.08	581.59	767.28
5	582.02	117.97	271.61	400.26	533.17	584.16	755.60
6	757.91	118.34	271.01	406.01	540.34	577.05	752.09

Figura 14 Tiempo promedio descompresión para un Sistema Alain sobre un sistema anfitrión, utilizando Slackware 12

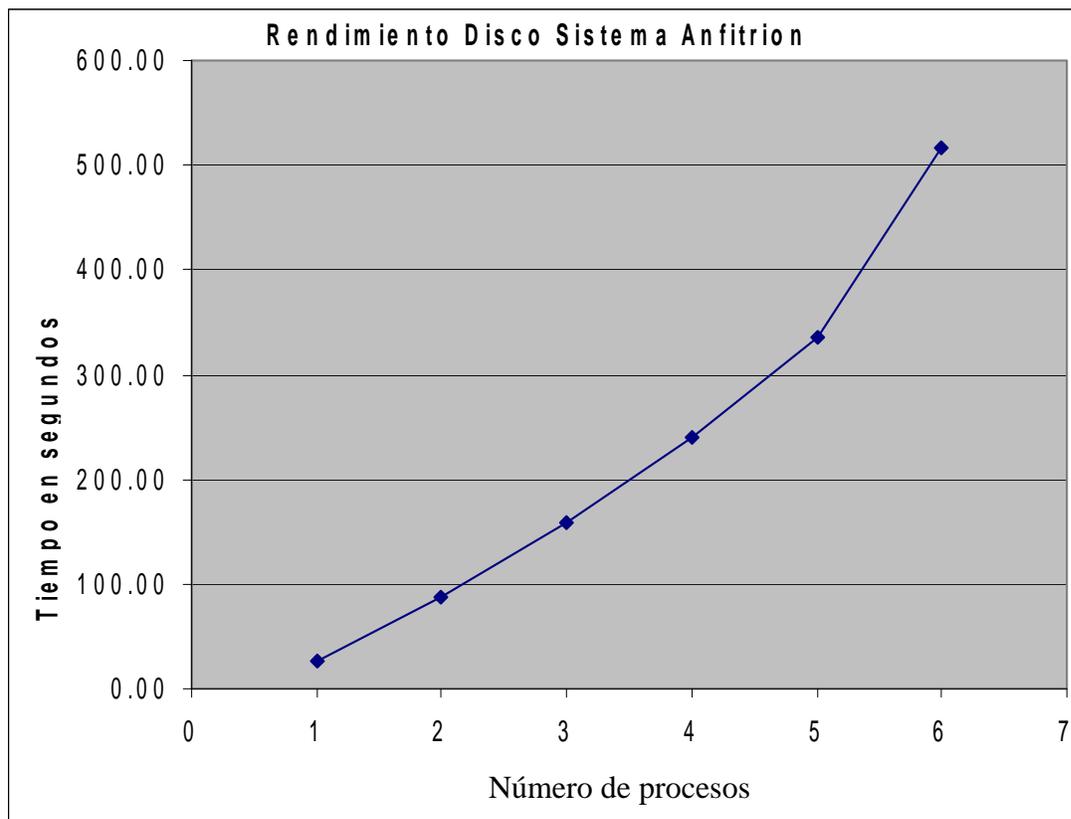


4.1.3.3 Resultados Sistema Alain sobre un sistema anfitrión, utilizando Windows XP SP2

Tabla XII. Rendimiento del Sistema Alain sobre Windows XP SP 2

# procesos	tiempo promedio en segundos	Tiempo en seg para N procesos					
		1	2	3	4	5	6
1	26.79	27.68	90.91	166.5	246.8	322.1	528.1
2	87.86	25.05	85.68	164.8	239.6	340.5	519
3	159.66	25.79	89.76	160.6	231.9	344.3	503.4
4	239.72	29.11	86.92	152.4	240.1	323.5	536.4
5	334.73	27.85	84.4	158.7	235.3	335.7	514.3
6	517.60	25.28	89.5	155	244.8	342.3	504.4

Figura 15 Tiempo promedio descompresión para un Sistema Alain sobre un sistema anfitrión, utilizando Windows XP SP2



4.1.3.4 Diferencia de rendimiento de acceso a disco entre el Sistema Alain y el sistema anfitrión

Tabla XIII. Diferencia de rendimiento de disco del Sistema Alain y el sistema anfitrión

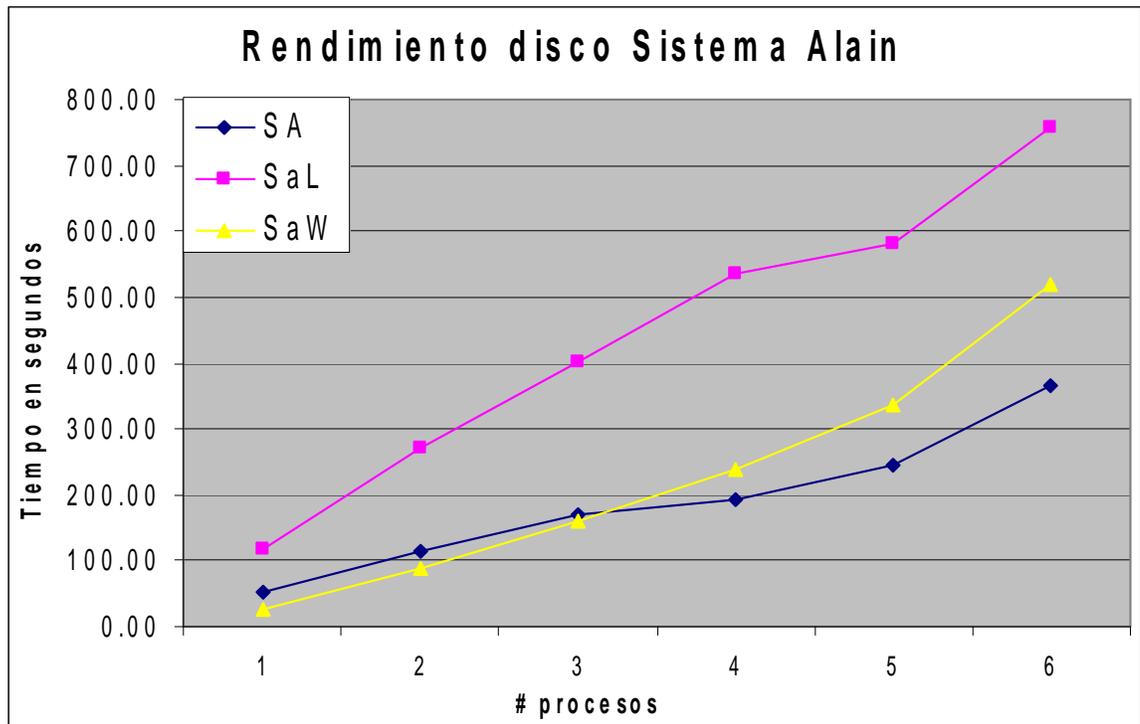
Tiempo en segundos para distintas configuraciones			
# procesos	SA	SaL	SaW
1	51.09	118.64	26.79
2	112.89	271.80	87.86
3	168.91	402.07	159.66
4	193.98	535.21	239.72
5	245.79	582.02	334.73
6	365.01	757.91	517.60

SA= Sistema anfitrión corriendo Slackware Linux 12

SaL= Sistema Alain en Sistema Anfitrión Slackware Linux 12

SaW= Sistema Alain en Sistema Anfitrión Windows XP

Figura 16 Diferencia de rendimiento de acceso a disco entre el Sistema Alain y el sistema anfitrión



4.1.4 Predicción del rendimiento de disco para los distintos sistemas

Con los datos adquiridos a través de las mediciones de tiempo del cálculo de descompresión de los fuentes del kernel en los distintos sistemas (anfitrión y Alain) se ha calculado haciendo uso de regresión lineal, regresión exponencial y regresión de potencias, curvas que describen el rendimiento de los sistemas y lo que es mas importante la diferencia de rendimiento que supone la virtualización para el Sistema Alain.

La penalización a la que se ve expuesta el Sistema Alain sobre un sistema anfitrión utilizando Slackware Linux 12 viene dada por:

Ecuación 4-6 Fórmula para el porcentaje de penalización CPU para Sistema Alain sobre un sistema anfitrión, utilizando Slackware Linux 12

$$y = \frac{18.593 + 121.719x}{e^{3.99183776X^{0.211923271}}}$$

Donde Y es el porcentaje de penalización que sufrirá el sistema y

X es el número de procesos ejecutándose simultáneamente

La penalización a la que se ve expuesta el Sistema Alain sobre un sistema anfitrión utilizando Windows XP SP2 viene dada por:

Ecuación 4-7 Fórmula para el porcentaje de penalización CPU para Sistema Alain sobre un sistema anfitrión utilizando Windows XP SP 2

$$y = \frac{-99.75 + 93.56x}{e^{3.99183776X^{0.211923271}}}$$

Donde Y es el porcentaje de penalización que sufrirá el sistema y
X es el número de procesos ejecutándose simultáneamente

4.1.4.1 Sistema anfitrión

Tabla XIV. Rendimiento disco sistema anfitrión

a	b	x	Medido	y exp pow	%diff	y exp lineal	%diff	y lineal	%diff
		1	51.09	54.15	6%	66.56	23%	47.23	8%
3.99184	0.21192	2	112.89	101.85	11%	94.59	19%	104.18	8%
		3	168.91	154.21	10%	134.44	26%	161.13	5%
3.84650	0.35155	4	193.98	211.67	8%	191.08	2%	218.09	11%
		5	245.79	274.34	10%	271.58	9%	275.04	11%
-9.72533	56.95343	6	365.01	342.25	7%	385.98	5%	332.00	10%
		7		415.44		548.59		388.95	
		8		493.93		779.69		445.90	
		9		577.76		1108.15		502.86	
		10	750.58	666.96	13%	1574.98	52%	559.81	34%

Correlación modelo exponencial potencia	0.99098
Covarianza modelo exponencial potencia	0.09305
Correlación modelo exponencial	0.96358
Covarianza modelo exponencial	1.23043
Correlación modelo lineal	0.97721
Covarianza modelo lineal	199.33700

La tabla muestra los valores para los distintos modelos para estimar el tiempo de descompresión, se muestra en negrita los tiempos más aproximados que en este caso pertenecen a una regresión de tipo potencia sobre datos a los que se aplicó el modelo exponencial (se les aplicó Logaritmo Natural).

El tiempo que le lleva a **X** procesos en paralelo dentro de un solo sistema Anfitrión completar la tarea de descompresión de los fuentes del kernel de linux esta dado por la fórmula

Ecuación 4-8 Fórmula de tiempo para la descompresión del kernel de Linux en función del numero de procesos en paralelo para un Sistema Alain sobre un sistema anfitrión utilizando Slackware 12

$$y = e^{3.99183776X^{0.211923271}}$$

Donde **Y** es el tiempo en segundos que tomara terminar el proceso y
X es el número de procesos ejecutándose simultáneamente

Donde **y** es el tiempo que le tomará al Sistema Alain llevar a cabo los **X** procesos paralelamente.

4.1.4.2 Sistema Alain sobre un sistema anfitrión utilizando Slackware 12

Tabla XV. Rendimiento Sistema Alain sobre un sistema anfitrión, utilizando Slackware 12

a	b	x	Medido	y exp pow	%diff	y exp lineal	%diff	y lineal	%diff	% overhead
		1	118.64	116.59	2%	163.53	27%	140.31	15%	-159%
4.759	0.178	2	271.80	217.81	25%	229.38	18%	262.03	4%	-157%
		3	402.07	325.90	23%	321.74	25%	383.75	5%	-149%
4.759	0.338	4	535.21	441.72	21%	451.29	19%	505.47	6%	-139%
		5	582.02	565.38	3%	633.00	8%	627.19	7%	-129%
18.593	121.719	6	757.91	696.86	9%	887.87	15%	748.90	1%	-119%
		7		836.08		1245.37		870.62		-110%
		8		982.99		1746.82		992.34		-101%
		9		1137.53		2450.18		1,114.06		-93%
		10	1055.00	1299.64	19%	3436.74	69%	1,235.78	15%	-85%

Correlación modelo exponencial potencia	0.99061
Covarianza modelo exponencial potencia	0.07817
Correlación modelo exponencial	0.94444
Covarianza modelo exponencial	1.18427
Correlación modelo lineal	
Covarianza modelo lineal	

La tabla muestra los valores para los distintos modelos para estimar el tiempo de descompresión, se muestra en negrita los tiempos mas aproximados que en este caso pertenecen a la regresión lineal.

El tiempo que le lleva a **X** procesos en paralelo dentro de un solo Sistema Alain sobre un sistema anfitrión utilizando Slackware 12 completar la tarea de descompresión de los fuentes del kernel de Linux esta dado por la fórmula

Ecuación 4-9 Fórmula de tiempo para cálculo de PI en función del número de procesos en paralelo para un Sistema Alain, sobre un sistema anfitrión utilizando Slackware 12

$$y = 18.593 + 121.719x$$

Donde Y es el tiempo en segundos que tomara terminar el proceso y
X es el número de procesos ejecutándose simultáneamente

Donde **y** es el tiempo que le tomara al Sistema Alain llevar a cabo los **X** procesos paralelamente

4.1.4.3 Sistema Alain sobre un sistema anfitrión, utilizando Windows XP SP 2

Tabla XVI. Rendimiento de un Sistema Alain sobre un sistema anfitrión, utilizando Windows XP SP 2

a	b	x	Medido	y exp pow	%diff	y exp lineal	%diff	y lineal	%diff	% over head
		1	26.79	28.76	7%	40.04	33%	(6.18)	533%	111%
3.36	0.35	2	87.86	73.06	20%	69.35	27%	87.38	1%	14%
		3	159.66	141.52	13%	120.12	33%	180.94	12%	-17%
3.14	0.55	4	239.72	240.44	0%	208.05	15%	274.51	13%	-30%
		5	334.73	377.05	11%	360.35	7%	368.07	9%	-34%
-99.75	93.56	6	517.60	559.61	8%	624.13	17%	461.64	12%	-35%
		7		797.53		1,081.00		555.20		-34%
		8		1101.49		1,872.31		648.76		-31%
		9		1483.50		3,242.87		742.33		-28%
		10	876.00	1957.07	55%	5,616.70	84%	835.89	5%	-25%

Correlación modelo exponencial potencia	0.99278
Covarianza modelo exponencial potencia	0.15519
Correlación modelo exponencial	0.97793
Covarianza modelo exponencial	327.472
Correlación modelo lineal	0.96765
Covarianza modelo lineal	1.92250

La tabla muestra los valores para los distintos modelos para estimar el tiempo de descompresión, se muestra en negrita los tiempos más aproximados que en este caso pertenecen a la regresión lineal.

El tiempo que le lleva a **X** procesos en paralelo dentro de un solo Sistema Alain sobre un sistema anfitrión utilizando Slackware 12 completar la tarea de descompresión de los fuentes del kernel de Linux esta dado por la fórmula

Ecuación 4-10 Fórmula de tiempo para cálculo de PI en función del número de procesos en paralelo para un Sistema Alain, sobre un sistema anfitrión utilizando Slackware 12

$$y = -99.75 + 93.56x$$

Donde Y es el tiempo en segundos que tomara terminar el proceso y
X es el número de procesos ejecutándose simultáneamente

Donde **y** es el tiempo que le tomara al Sistema Alain llevar a cabo los **X** procesos paralelamente

CONCLUSIONES

1. Crecimiento lineal del tiempo con respecto al número de procesos que corren simultáneamente en un Sistema Alain para la prueba de rendimiento de CPU, esto es un excelente indicador que el overhead al que está expuesto el sistema para aplicaciones intensivas de CPU es realmente poco, volviendo la plataforma virtual una excelente opción para aplicaciones que hacen uso del CPU de un modo intensivo como lo pueden ser granjas de renderización de animación tridimensional
2. Tiempos de la prueba de CPU para el Sistema Alain cercanos al del sistema anfitrión corriendo Slackware 12 y Windows XP SP 2, esto es un excelente indicador del rendimiento del Sistema Alain, ya que para soluciones que se planeen implementar en sistemas que demandan alto rendimiento, los beneficios que ofrece la infraestructura virtual serán de aplastante peso para la selección de una solución de este tipo.
3. La virtualización representa una penalización menor para el Sistema Alain al estar corriendo sobre un sistema anfitrión que hace uso de Windows XP SP2, debido a la menor penalización de acceso a disco.
4. La penalización de CPU a la que se ve expuesto el Sistema Alain en

una máquina virtual corriendo en un sistema anfitrión, utilizando Slackware Linux 12 en comparación al mismo proceso corriendo en

la máquina física viene dado por: $y = \frac{0.7626667 + 13.264x}{0.733333 + 13.273x}$

donde Y es el porcentaje de penalización que sufrirá el sistema y X es el número de procesos ejecutándose simultáneamente

5. La penalización de CPU a la que se ve expuesto el Sistema Alain en una máquina virtual corriendo en un sistema anfitrión utilizando Windows XP SP2 en comparación al mismo proceso corriendo en la

máquina física viene dado por: $y = \frac{2.442667 + 12.80828571x}{0.733333 + 13.273x}$

donde Y es el porcentaje de penalización que sufrirá el sistema y X es el número de procesos ejecutándose simultáneamente

6. La penalización de disco a la que se ve expuesto el Sistema Alain en una máquina virtual corriendo en un sistema anfitrión, utilizando Slackware Linux 12 en comparación al mismo proceso corriendo en

la máquina física viene dado por: $y = \frac{18.593 + 121.719x}{e^{3.99183776X^{0.211923271}}}$

donde Y es el porcentaje de penalización que sufrirá el sistema y X es el número de procesos ejecutándose simultáneamente

7. La penalización de disco a la que se ve expuesto el Sistema Alain en una máquina virtual corriendo en un sistema anfitrión, utilizando Windows XP SP 2 en comparación al mismo proceso corriendo en la

máquina física viene dado por: $y = \frac{-99.75 + 93.56x}{e^{3.99183776X^{0.211923271}}}$ donde Y

es el porcentaje de penalización que sufrirá el sistema y X es el número de procesos ejecutándose simultáneamente

RECOMENDACIONES

1. Previa implementación de un sistema en un contenedor virtual, se debe estudiar los tiempos de respuesta necesarios para el sistema y estimar la penalización a la que se ve expuesto el sistema, para realizar las previsiones necesarias.
2. Para sistemas que harán uso extensivo del disco y dependen directamente del rendimiento de este, se debe evaluar cuidadosamente la implementación en contenedores virtuales debido a la penalización del acceso a disco que estos sufren.
3. Los contenedores virtuales son una excelente alternativa para sistemas que hacen uso extensivo del CPU debido a la poca penalización a la que se ven expuestos.
4. Dentro de los productos de contenedores virtuales VMWare Server es un producto que destaca por su costo gratuito, implementación de tecnologías de punta, rendimiento, estabilidad, capacidades de migración y soporte.

REFERENCIAS ELECTRÓNICAS

1. VMware (2007), White papers on server consolidation, blade servers, software testing and more
<<http://www.vmware.com/solutions/whitepapers.html>> (12 septiembre 2007)
2. VMware (2007), Intro to Virtualization, Virtual Machine, Server Consolidation <<http://www.vmware.com/virtualization/>> (12 septiembre 2007)
3. VMware (2007), Why Virtualize Virtualization Improves Performance and Efficiency <<http://www.vmware.com/overview/why.html>> (12 septiembre 2007)
4. VMware (2007), Virtualization History
<<http://www.vmware.com/overview/history.html>> (12 septiembre 2007)
5. VMware (2007), Virtual Machines
<<http://www.vmware.com/overview/vmachine.html>> (12 septiembre de 2007)
6. VMware (2007), virtual Infraestructure Overview, Virtualization, Virtual

Servers <<http://www.vmware.com/vinfrastructure/>> (12 septiembre de 2007)

7. VMware (2007), Resources

<<http://www.vmware.com/overview/performance/resources.html>> (12 septiembre 2007)

8. VMware (2007), White papers on server consolidation, blade servers, software testing and more
<<http://www.vmware.com/overview/performance/resources.html>> (12 septiembre 2007)
9. Wikipedia (2007), Virtualization
<<http://en.wikipedia.org/wiki/Virtualization>> (12 septiembre 2007)
10. Wikipedia (2007), Hypervisor <<http://en.wikipedia.org/wiki/Hypervisor>> (12 septiembre 2007)
11. Wikipedia (2007), x86 virtualization
<http://en.wikipedia.org/wiki/X86_virtualization> (12 septiembre 2007)
12. Wikipedia (2007), VMware <<http://en.wikipedia.org/wiki/VMware>> (12 septiembre 2007)
13. Wikipedia (2007), Paravirtualization
<<http://en.wikipedia.org/wiki/Paravirtualization>> (12 septiembre 2007)
14. Wikipedia (2007), Comparison of virtual machines
<http://en.wikipedia.org/wiki/Comparison_of_virtual_machines> (12 septiembre 2007)

15. VMware (2007) VMotion Overview
<<http://www.vmware.com/products/vi/vc/vmotion.html>> (30 septiembre 2007)
16. VMware (2007) DRS Overview <
<http://www.vmware.com/products/vi/vc/drs.html>> (30 septiembre 2007)
17. Wikipedia (2007) Linux <<http://en.wikipedia.org/wiki/Linux>> (4 octubre 2007)
18. Wikipedia (2007) Linux Distribution
<http://en.wikipedia.org/wiki/Linux_distribution> (4 de octubre 2007)
19. TTY Linux (2007) homepage
<<http://www.minimalinux.org/ttylinux/showpage.php?pid=1>> (6 octubre 2007)
20. TTY Linux (2007) Downloads
<<http://www.minimalinux.org/ttylinux/showpage.php?pid=4>> (7 de octubre 2007)
21. Linux Kernel (2007) Archive < <http://www.kernel.org/>> (7 de octubre 2007)
22. SYSLINUX (2007) The easy to use Linux bootloader
<<http://syslinux.zytor.com/>> (10 octubre 2007)

23. Elephant's Homepage (2007) How to calculate PI

<<http://elephant.linux.net.cn/articles/pi.php>> (10 octobre 2007)

BIBLIOGRAFÍA

1. Scout Rixner “Network Virtualization, breaking the performance barrier” ACM queue (Estados Unidos) (Volumen 6 numero 1): páginas 37-45. Enero 2008.
2. Ulrich Drepper “The cost of virtualization” ACM queue (Estados Unidos) (Volumen 6 numero 1): páginas 30-35. Enero 2008.
3. VMware (2007) Virtualization White paper
<<http://www.vmware.com/pdf/virtualization.pdf>> (21 de septiembre 2007).

